

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:42:25 ; Search time 18.3444 Seconds
(without alignments) 287.144 Million cell updates/sec

Title: US-09-834-794a-1

Perfect score: 677
Sequence: 1 MOORGIAIYALVACALHAS.....RNSNRAHQKHEFYGHKTPY 127

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	99.4	127	1 SY28_HUMAN	O99113 homo sapien
2	418.5	61.8	130	1 SY28_MOUSE	O91112 mus musculu
3	162.5	24.0	120	1 SY27_MOUSE	O92110 mus musculu
4	147.5	21.8	112	1 SY27_HUMAN	O94443 homo sapien
5	94	13.9	150	1 SY25_HUMAN	O15444 homo sapien
6	92	13.6	144	1 SY25_MOUSE	O35903 mus musculu
7	78	11.5	119	1 SY24_MOUSE	O91110 mus musculu
8	76	11.2	91	1 SY05_BOVIN	O91919 bos taurus
9	75	11.1	97	1 SY20_MOUSE	O89093 mus musculu
10	74.5	11.0	96	1 SY20_RAT	O97884 rattus norv
11	71.5	10.6	90	1 SY04_CHICK	O90826 gallus gall
12	70	10.3	2056	1 CBP1_CAEEL	P34545 caenorhabdi
13	68	10.0	173	1 Y19A_ECOLI	O47309 escherichia
14	67	9.9	901	1 DSC2_HUMAN	O02427 homo sapien
15	66	9.7	91	1 SY05_HUMAN	P13501 homo sapien
16	65.5	9.7	97	1 BOTA_RAT	P97545 rattus norv
17	65	9.6	120	1 SY16_HUMAN	O15467 h small ind
18	65	9.6	146	1 Y220_HUMAN	O92617 homo sapien
19	64.5	9.5	91	1 SY05_CAVPO	P97272 cavia porce
20	64.5	9.5	187	1 HPAP_RALSO	P35651 talstonia s
21	64.5	9.5	1361	1 GLI4_XENLA	O91661 xenopus lae
22	64	9.5	727	1 IF2M_HUMAN	P46149 homo sapien
23	63.5	9.4	99	1 SY08_HUMAN	P80075 homo sapien
24	63.5	9.4	109	1 SZ13_HUMAN	O43937 homo sapien
25	63.5	9.4	289	1 PHIC_LISMO	P33378 listeria mo
26	63.5	9.4	295	1 XERC_HAEIN	P44818 haemophilus
27	63.5	9.4	719	1 YP62_CAEEL	O09437 caenorhabdi
28	63	9.3	94	1 SY26_HUMAN	O92258 homo sapien
29	63	9.3	285	1 RL2_MYCCE	P47400 mycoplasma
30	63	9.3	663	1 GLI3_CHICK	P55879 gallus gall
31	62.5	9.2	97	1 SY08_MOUSE	O92121 mus musculu
32	62.5	9.2	930	1 AD19_MOUSE	O35614 mus musculu
33	62	9.2	96	1 SY20_HUMAN	P78556 homo sapien

34	62	9.2	126	1 SZ09_MOUSE	P18340 mus musculu
35	62	9.2	385	1 PO3B_XENLA	P70030 xenopus lae
36	62	9.2	428	1 FXB2_MOUSE	O64733 mus musculu
37	62	9.2	534	1 GCR2_YEAST	O01722 saccharomyc
38	62	9.2	560	1 EAA5_HUMAN	O00341 homo sapien
39	61.5	9.1	97	1 IL8_HORSE	O62812 equus caball
40	61.5	9.1	495	1 TRME_TREPA	O83561 treponema p
41	61.5	9.1	1167	1 WC1_NEUCR	O01371 neurospora
42	61.5	9.1	1411	1 YK63_CAEEL	P34342 caenorhabdi
43	61	9.0	325	1 NPDL_HUMAN	O99455 homo sapien
44	61	9.0	382	1 PO3A_XENLA	P31365 xenopus lae
45	61	9.0	567	1 TGR2_RAT	P38438 rattus norv

ALIGNMENTS

RESULT 1
SY28_HUMAN STANDARD; PRT; 127 AA.
ID SY28_HUMAN
AC O9NRJ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine A28 precursor (CCL28) (Mucosae-associated epithelial chemokine) (MCC) (CC1 protein).
GN SCYA28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE-Fetal heart, and osteoblast;
RC MEDLINE=20357357; PubMed=10761587;
RX MEDLINE=20357357; PubMed=10761587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Honey B., Catron D., Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J., Kershenovich D., Smith K., McLanahan T., Vicari A.P., Zlotnik A.;
RT "Identification of a novel chemokine (CCL28), which binds CCR10 (GPR2)."
RT J. Biol. Chem. 275:22313-22323 (2000).
RL [2]
RN SEQUENCE FROM N.A. AND RECEPTOR INTERACTION.
RP MEDLINE=20432268; PubMed=10975800;
RX Pan J., Kunkel E.J., Gossier U., Lazarus N., Landon P., Broadwell K., Vierra M.A., Genovese M.C., Butcher E.C., Soler D.;
RA "A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial cells in mucosal tissues."
RT J. Immunol. 165:2943-2949 (2000).
RL [3]
RN SEQUENCE FROM N.A.
RP Zhang W., He L., Yuan Z., Wan T., Cao X.;
RT "A novel CC chemokine homology with TECK."
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CHEMOTACTIC ACTIVITY FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS. BINDS TO CCR3 AND CCR10 AND INDUCES CALCIUM MOBILIZATION IN A DOSE-DEPENDENT MANNER.
CC -1- TISSUE SPECIFICITY: SECRETED.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED BY EPITHELIAL CELLS OF DIVERSE TISSUES INCLUDING NORMAL AND PATHOLOGICAL COLON, SALIVARY GLAND, MAMMARY GLAND, TRACHEA AND RECTUM. ALSO FOUND IN PROSTATE, SPLEEN, THYROID, PSORIASIS SKIN AND IN LOWER LEVELS IN PERIPHERAL BLOOD LEUCOCYTES, SMALL INTESTINE, PETER'S PATCHES, STOMACH AND NORMAL SKIN.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE C-C) (CHEMOKINE CC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@sb-sib.ch.
CC -----
DR EMBL: AF220210; AAF87205.1;
DR EMBL: AF266504; AAG16691.1;
DR EMBL: AF110384; AAG43193.1;
DR MIM: 605240;
DR InterPro: IPR000827; CC_chemokine_sml.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
KW Cytokine; Chemotaxis; Signal; Glycoprotein.
FT SIGNAL 1 19
FT CHAIN 20 127
FT DISULFID 30 58
FT DISULFID 31 73
FT CARBOHYD 78 78
FT CARBOHYD 78 78
SQ SEQUENCE 127 AA; 14280 MW; 3E8551A63A2C8D62 CRC64;

Query Match 99.4%; Score 673; DB 1; Length 127;
Best Local Similarity 98.4%; Pred. No. 8.4e-68;
Matches 125; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 MOORGLAIYALAVCALHASEALPIASSCCTFEVSHHISRLERVMCRIOADGDCDL 60
DB 1 MOORGLAIYALAVCALHASEALPIASSCCTFEVSHHISRLERVMCRIOADGDCDL 60
OY 61 AAVILHVKKRRLICVSPHNHTVQMKVQAAXKNGKGVCHRRKHGKRSNRAHQKHET 120
DB 61 AAVILHVKKRRLICVSPHNHTVQMKVQAAXKNGKGVCHRRKHGKRSNRAHQKHET 120
OY 121 YGHKTPY 127
DB 121 YGHKTPY 127

RESULT 2
SY28_MOUSE
ID SY28_MOUSE STANDARD: PRT: 130 AA.
AC Q9JIL2;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Small inducible cytokine A28 precursor (CCL28).
GN SCYA28.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RC TISSUE-Kidney;
RX MEDLINE=20557357; PubMed=10781587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Homey B., Catron D.,
RA Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Adams J.,
RA Kershenevich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.;
RT "Identification of a novel chemokine (CCL28), which binds CCR10
RT (GPR2).";
RL J. Biol. Chem. 275:22313-22323(2000).
CC -1- FUNCTION: CHEMOKINE FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS
CC (BY SIMILARITY). BINDS TO CCR10 AND INDUCES CALCIUM MOBILIZATION
CC IN A DOSE-DEPENDENT MANNER.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN TESTIS, EPITHELIAL CELLS
CC OF NORMAL COLON, KIDNEY, PEYER'S PATCHES, LYMPH NODES. ALSO FOUND
CC IN LOWER LEVELS IN BRAIN, SPLEEN AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE INTERCINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: AF220238; AAF87206.1;
DR MGD: MGI:1861731; SCYA28.
DR InterPro: IPR000827; CC_chemokine_sml.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00199; SCY; 1.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
KW Cytokine; Chemotaxis; Signal.
FT SIGNAL 1 16
FT CHAIN 17 130
FT DISULFID 30 58
FT DISULFID 31 73
FT CARBOHYD 78 78
FT CARBOHYD 78 78
SQ SEQUENCE 130 AA; 14570 MW; 6F3B909A4F97F013 CRC64;

Query Match 61.8%; Score 418.5; DB 1; Length 130;
Best Local Similarity 63.2%; Pred. No. 1.5e-39;
Matches 79; Conservative 14; Mismatches 27; Indels 5; Gaps 1;

OY 1 MOORGLAIYALAVCALHASEALPIASSCCTFEVSHHISRLERVMCRIOADGDCDL 60
DB 1 MOORGLAIYALAVCALHASEALPIASSCCTFEVSHHISRLERVMCRIOADGDCDL 60
OY 61 AAVILHVKKRRLICVSPHNHTVQMKVQAAXKNGKGVCHRRKHGKRSNRAHQKHET 120
DB 61 AAVILHVKKRRLICVSPHNHTVQMKVQAAXKNGKGVCHRRKHGKRSNRAHQKHET 120
OY 121 YGHKTPY 125
DB 121 YGHKTPY 125

RESULT 3
SY27_MOUSE
ID SY27_MOUSE STANDARD: PRT: 120 AA.
AC Q9Z1X0;
DT 30-MAY-2000 (Rel. 39; Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Small inducible cytokine A27 precursor (CCL27) (CC chemokine IL2) (IL-
DE 11 RALPHA-locus chemokine) (MILC) (ALP) (Skinline) (ESKine)
DE (Cutaneous T-cell attracting chemokine) (CTACK).
GN SCYA27 OR IL2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99262167; PubMed=10329455;
RA Thomas R., Broxmeyer H.E., Kim C., Christopherson K. II, Hou Y.-H.;
RT "Isolation of ALP, a novel divergent murine CC chemokine with a unique
RT carboxy terminal extension.";
RL Biochem. Biophys. Res. Commun. 258:737-740(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2006096; PubMed=10556532;
RA Ishikawa-Mochizuki I., Kitaara M., Baba M., Nakayama T., Izawa D.,
RA Imai T., Yamada H., Hieshima K., Suzuki R., Nomiyama H., Yoshie O.;
RT "Molecular cloning of a novel CC chemokine, Interleukin-11 receptor
RT alpha-locus chemokine (ILC), which is located on chromosome 9p13 and a
RT potential homologue of a CC chemokine encoded by molluscan contagiosum
RT virus.";
RL FEBS Lett. 460:544-548(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056268; PubMed=10588729;
RA Morales J., Homey B., Vicari A.P., Hudek S., Oldham E., Hedrick J.,
RA Orocco R., Copeland N.G., Jenkins N.A., McEvoy L.M., Zlotnik A.;
RT "CTACK, a skin-associated chemokine that preferentially attracts
RT skin-homing memory T cells.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 96:14470-14475(1999).
CC -1- FUNCTION: CHEMOTACTIC FACTOR THAT ATTRACTS SKIN-ASSOCIATED MEMORY
CC T-LYMPHOCYTES. MAY PLAY A ROLE IN MEDIATING HOMING OF LYMPHOCYTES
CC TO CUTANEOUS SITES. MAY PLAY A ROLE IN CELL MIGRATION DURING
CC EMBRYOGENESIS.
CC -1- TISSUE SPECIFICITY: TESTIS, HEART AND LIVER AND SKIN. EXPRESSED AT
CC LOWER LEVELS IN THE KIDNEY. ALSO FOUND IN THE PLACENTA.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-2 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF099931; AAD04163.1; -.
DR EMBL: AB013398; BAA8474.1; -.
DR EMBL: AF082392; AAD41237.1; -.
DR MGI: 1343459; Scy27.
DR InterPro: IPR000827; CC_chemkine_sml.
DR SMART: SM001811; Chemokine_IL8.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
KM CYTOKINE; Signal.
FT CHAIN 1 25 POTENTIAL.
FT DISULFID 26 120 SMALL INDUCIBLE CYTOKINE A27.
FT DISULFID 34 63 BY SIMILARITY.
FT DISULFID 35 78 BY SIMILARITY.
SQ SEQUENCE 120 AA; 13441 MW; D1EEI1270AB020BF CRC64;

Query Match 24.0%; Score 162.5; DB 1; Length 120;
Best Local Similarity 41.8%; Pred. No. 3.5e-11;
Matches 33; Conservative 14; Mismatches 29; Indels 3; Gaps 2;

OY 19 ASELILPIAS--SCCTEYSHH-ISRLLERYVMCRIPADDDCDLAAVILHVKRKRICVS 75
DB 21 APEALPLPSSCTCTOLYRPLSRLLRKYVHMEIDADDDCHLQAVYHLARSCVH 80
OY 76 PHNFTVAKMKVQAAXKNG 94
DB 81 PGNRSIARWLERGGKRLQG 99

RESULT 4
SY27_HUMAN STANDARD; PRT; 112 AA.
ID SY27_HUMAN STANDARD; PRT; 112 AA.
AC G9Y4X3;
DT 30-MAY-2000 (Rel. 39, Last Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine A27 precursor (CCP27) (CC chemokine IL8) (IL-
DE 11 Ralphi-j locus chemokine) (Skinkine) (Eskine) (Cutaneous T-cell
DE attracting chemokine) (CRACK).
OS SCY27 OR IL8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-30.
RC TISSUE=Thymus;
RA MEDLINE=20026096; PubMed=10556532;
RA Ishikawa-Mochizuki I., Kikawa M., Baba M., Nakayama T., Izawa D.,
RA Inai T., Yamada T., Hieshima K., Suzuki R., Nomiyama H., Yoshie O.,
RA "Molecular cloning of a novel CC chemokine, Interleukin-11 receptor
RA alpha-j locus chemokine (IL8), which is located on chromosome 9p13 and a
RA potential homologue of a CC chemokine encoded by molluscum contagiosum,
RA Virus."

```

```

RL FEBS Lett. 460:544-548(1999).
RN [2]
RP SEQUENCE FROM N.A.
EX MEDLINE=20056268; PubMed=10588729;
RA Morales-J., Homey B., Vicari A.P., Hudak S., Oldham E., Hedrick J.,
RA Orozco R., Copeland N.G., Jenkins N.A., McEvoy L.M., Zlotnik A.,
RA "CRACK, a skin-associated chemokine that preferentially attracts
RA skin-homing memory T cells."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14470-14475(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Zaballios A., Gutierrez J., Marquez G., Hromas R.,
RA "CCL27, the human homologue of murine ALP chemokine."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP RECEPTOR INTERACTION.
RX MEDLINE=20191998; PubMed=10725697;
RA Homey B., Wang W., Soto H., Buchanan M.E., Wiesenborn A., Catron D.,
RA Muller A., McClanahan T.K., Dieu-Nosjean M.C., Orozco R., Ruzicka T.,
RA Lehmann P., Oldham E., Zlotnik A.;
RA "Cutting edge: the orphan chemokine receptor G protein-coupled
RA receptor-2 (GPR-2, CCR10) binds the skin-associated chemokine CCL27."
RL J. Immunol. 164:3465-3470(2000).
CC -1- FUNCTION: CHEMOTACTIC FACTOR THAT ATTRACTS SKIN-ASSOCIATED MEMORY
CC T-LYMPHOCYTES. MAY PLAY A ROLE IN MEDIATING HOMING OF LYMPHOCYTES
CC TO CUTANEOUS SITES. BINDS TO CCR10.
CC -1- TISSUE SPECIFICITY: TESTIS, THYMUS, PLACENTA, OVARY AND SKIN.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB010445; BAA87046.1; -.
DR EMBL: AF082393; AAD41238.1; -.
DR EMBL: AJ243542; CAB46983.1; -.
DR Genew; HGNC:10626; SCY27.
DR MIM: 604833; -.
DR InterPro: IPR000827; CC_chemkine_sml.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00189; SCY; 1.
DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
KM CYTOKINE; Signal.
FT CHAIN 1 24 SMALL INDUCIBLE CYTOKINE A27.
FT DISULFID 25 112 BY SIMILARITY.
FT DISULFID 33 62 BY SIMILARITY.
FT DISULFID 34 77 BY SIMILARITY.
SQ SEQUENCE 112 AA; 12618 MW; 7B5940E147AFB820 CRC64;

Query Match 21.8%; Score 147.5; DB 1; Length 112;
Best Local Similarity 37.0%; Pred. No. 1.5e-09;
Matches 27; Conservative 16; Mismatches 27; Indels 1; Gaps 1;

OY 23 ILPIASCTEYSHH-ISRLLERYVMCRIPADDDCDLAAVILHVKRKRICVSPHNHTV 81
DB 26 ILPISTACTCTOLYRPLSRLLRKYVHMEIDADDDCHLQAVYHLARSCVH 85
OY 82 KQMKVQAAXKNG 94
DB 86 SQWFEHQRKLHG 98

RESULT 5
SY25_HUMAN STANDARD; PRT; 150 AA.
ID SY25_HUMAN STANDARD; PRT; 150 AA.
AC O15444; O96KJ7;
DT 15-JUL-1998 (Rel. 36, Created)

```


Matches 26; Conservative 15; Mismatches 37; Indels 12; Gaps 5;

OY 1 MOORGLAVALAVCALHASEALIPASS---CCTEVSHHSRLRLRYVMNCRIQADGD 57
 DB 1 MKVSAFAVFLMAALCALPASPASPYASDTTPCCFA---YISNP-LPRTHVQGYFTSSK 56
 OY 58 CDLAIVILHVKRR-ICVSPHNTVQWKK 86
 DB 57 GSAVAVFTRKKROYCANPE---KKWVR 82

RESULT 9

SY20_MOUSE STANDARD; PRT: 97 AA.

ID SY20_MOUSE 089093; Q921X3;
 AC 089093; Q921X3; Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small inducible cytokine A20 precursor (CCL20) (Macrophage
 inflammatory protein 3 alpha) (MIP-3-alpha) (Liver and activation-
 regulated chemokine) (CC chemokine LARC) (Beta chemokine exodus-1) (CC
 chemokine ST38).
 DE GN SCYA20 OR LARC.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RX MEDLINE=99113931; PubMed=9916893;
 RA utans-schneitz U., Lorez H., Klinkert W.E.F., da Silva J.,
 RA Lesslauer W.;
 RT "A novel rat CC chemokine, identified by targeted differential
 RT display, is upregulated in brain inflammation.";
 RL J. Neuroimmunol. 92:179-190(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RA Villares R.;
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A. (SHORT ISOFORM).
 RC STRAIN-BALB/c;
 RA McCall S.R., Carney T., Neote K.S.;
 RT "Cloning and characterization of a murine homolog of macrophage
 RT inflammatory protein 3 alpha.";
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (LONG ISOFORM), AND CHARACTERIZATION.
 RC TISSUE=Thymus;
 RX MEDLINE=99077268; PubMed=9862452;
 RA Valcova R., Zaballos A., Gutierrez J., Martin P., Roncal F.,
 RA Alber J.P., Ardayin C., Marquez G.;
 RT "Molecular cloning, functional characterization and mRNA expression
 RT analysis of the murine chemokine receptor CCR6 and its specific ligand
 RT MIP-3alpha.";
 RL MIP-3alpha.";
 RL FEBS Lett. 440:188-194(1998).
 RN [5]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND SEQUENCE OF N-TERMINUS.
 RC TISSUE=Liver;
 RX MEDLINE=99161843; PubMed=10064080;
 RA Tanaka Y., Imai T., Baba M., Ishikawa I., Uehira M., Nomiya H.,
 RA Yoshie O.;
 RT "Selective expression of liver and activation-regulated chemokine
 RT (LARC) in intestinal epithelium in mice and humans.";
 RL Eur. J. Immunol. 29:633-642(1999).
 CC -1- FUNCTION: CHEMOTACTIC FACTOR THAT ATTRACTS LYMPHOCYTES AND,
 CC SLIGHTLY, NEUTROPHILS, BUT NOT MONOCYTES (BY SIMILARITY). MAY BE
 CC INVOLVED IN FORMATION AND FUNCTION OF THE MUCOSAL LYMPHOID TISSUES
 CC BY ATTRACTING LYMPHOCYTES AND DENDRITIC CELLS TOWARDS EPITHELIAL
 CC CELLS. BINDS TO CCR6.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PROMINENTLY EXPRESSED IN THE SMALL INTESTINE,
 CC COLON AND APPENDIX. ALSO FOUND IN THYMUS, SPLEEN, LYMPH NODE AND
 CC LUNG. THE LONG FORM MIGHT BE DOMINANT IN INTESTINAL, AND THE SHORT
 CC FORM IN LYMPHOID TISSUES.
 CC -1- INDUCTION: BY LIPOPOLYSACCHARIDE (LPS), TNF-ALPHA AND INTERLEUKIN-
 CC 1. IFN-GAMMA ALONE SHOWED NO EFFECT, BUT POTENTIATED THE EFFECT OF
 CC TNF.
 CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
 CC C-C) (CHEMOKINE CC).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----

DR EMBL; AF053313; AAC78295.1; -;
 DR EMBL; AJ007862; CAA07714.1; -;
 DR EMBL; AF099052; AAC78680.1; -;
 DR EMBL; AJ222694; CAA10948.1; -;
 DR EMBL; AB015136; BAA76955.1; -;
 DR HSP; P13236; 1HDM.
 DR MGD; MG1:1329031; Scya20.
 DR InterPro; IPR000827; CC_Chemokine_sml.
 DR InterPro; IPR001811; Chemokine_IL8.
 DR Pfam; PF00048; IL8; 1.
 DR SMART; SM00199; SCY; 1.
 DR PROSITE; PS00472; SMALL_CYTOKINES_CC; 1.
 DR Cytokine; Chemotaxis; Inflammatory response; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 27
 FT CHAIN 28 97 SMALL INDUCIBLE CYTOKINE A20.
 FT DISULFID 33 59 BY SIMILARITY.
 FT DISULFID 34 75 BY SIMILARITY.
 FT VARSPIC 28 28 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 97 AA; 10826 MW; 6E0C73C8AC80A80A CRC64;

Query Match 11.1%; Score 75; DB 1; Length 97;
 Best Local Similarity 34.1%; Pred. No. 0.14;
 Matches 28; Conservative 8; Mismatches 38; Indels 8; Gaps 5;

OY 6 LAIVALAVCALHASEALIPASSCCTEVSHHSRLRLRYVMNCRIQADGDCLAA 62
 DB 8 LFLALAVLVLHLCSQAASNYDCLS---YIGPLPSRAIVGTR-QMADECDINA 63
 OY 63 VILHV-KRRKICVSPHNTVQK 83
 DB 64 IIFTRKRSVCADPKQNWKR 85

RESULT 10

SY20_RAT STANDARD; PRT: 96 AA.

ID SY20_RAT P97884;
 AC P97884;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small inducible cytokine A20 precursor (CCL20) (Macrophage
 inflammatory protein 3 alpha) (MIP-3-alpha) (Liver and activation-
 regulated chemokine) (CC chemokine LARC) (Beta chemokine exodus-1) (CC
 chemokine ST38).
 DE GN SCYA20 OR ST38.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Kellner G.S., Maciejewski-Lenoir D., Lee E.D., Maki R.A.;
 RT "Activated rat astrocytes produce the CC chemokine Exodus.";

RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF N-TERMINUS.
 RC STRAIN-Fischer 344; TISSUE-Brain;
 RX MEDLINE-99113931; PubMed-9916893;
 RA Ucanas-Schneitz U., Lorez H., Klinkert W.E.F., da Silva J.,
 RA Lesslauer W.;
 RT "A novel rat CC chemokine, identified by targeted differential
 RT display, is upregulated in brain inflammation.";
 RL J. Neuroimmunol. 92:179-190(1998).
 CC -1- FUNCTION: CHEMOTACTIC FACTOR THAT ATTRACTS LYMPHOCYTES AND,
 CC SLIGHTLY, NEUTROPHILS, BUT NOT MONOCYTES (BY SIMILARITY). MAY PLAY
 CC A ROLE IN MODULATING INFLAMMATORY CELL RECRUITMENT TO THE CNS AND
 CC THEREFORE CONTRIBUTE TO TISSUE INJURY IN ISCHEMIC STROKE AND
 CC AUTOMUNE DISEASES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: LOW LEVELS IN THYMUS AND LUNG.
 CC -1- INDUCTION: BY TNF IN EXPERIMENTAL ALLERGIC PANENCAPALOMYELITIS
 CC (EAP) AND BY TNF AND IL-1 IN PRIMARY ASTROCYTES.
 CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
 CC C-C) (CHEMOKINE CC).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U90447; AAB61459.1; -;
 DR EMBL: AF053312; AAC78294.1; -;
 DR InterPro: IPR000827; CC-chemkine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR SMART: SM00048; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
 DR CycloLine: Chemotaxis; Inflammatory response; Signal.
 KW SIGNAL 1 26
 FT CHAIN 27 96 SMALL INDUCIBLE CYTOKINE A20.
 FT DISULFID 31 58 BY SIMILARITY.
 FT DISULFID 32 74 BY SIMILARITY.
 SQ SEQUENCE 96 AA: 10875 MW: 78323410F2B26DF CRC64;
 Query Match 11.0%; Score 74.5; DB 1; Length 96;
 Best Local Similarity 31.5%; Pred. No. 0.16; Indels 15; Gaps 6;
 Matches 29; Conservative 13; Mismatches 35;
 Oy 1 MOGRGLAIVALA-VCAALHASEALLPIASCC---TEVSHHSIRRLERVMKRIQADG 56
 Db 1 MACKRLPEPLAAGVLAIVLACSOSEASNFDCCLTYTKNVYHHRNFVGT---QMADE 56
 Oy 57 DCDLAAYLVHVKRRX-ICVSPHNHTVKO-WMK 86
 Db 57 ACDINAILFHLKRSKRVCAADP-----KOIWK 83
 RESULT 11
 SY04_CHICK STANDARD: PRT: 90 AA.
 AC Q90826; Q910C9;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small inducible cytokine A4 homolog precursor (Macrophage inflammatory
 DE protein 1-beta homolog).
 GN SCY4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-Bone marrow;
 RX MEDLINE-95369710; PubMed-7642115;
 RA Petrenko O., Ischenko I., Enrietto P.J.;
 RT "Isolation of a cDNA encoding a novel chicken chemokine homologous to
 RT mammalian macrophage inflammatory protein-1 beta.";
 RL Gene 160:305-306(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hughes S.M., Bunnstead N.;
 RT "Mapping of the gene encoding the chicken homologue of the mammalian
 RT chemokine SCY4.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 14-90 FROM N.A.
 RA Petrenko O., Enrietto P.J.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MONOKINE WITH INFLAMMATORY AND CHEMOKINETIC PROPERTIES
 CC (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
 CC C-C) (CHEMOKINE CC).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L34553; AAA48747.1; -;
 DR EMBL: AJ243034; CAB45103.1; -;
 DR HSSP: P13236; IHM.
 DR InterPro: IPR000827; CC-chemkine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; 1.
 DR CycloLine: Chemotaxis; Signal.
 KW SIGNAL 1 21
 FT CHAIN 22 90 SMALL INDUCIBLE CYTOKINE A4 HOMOLOG.
 FT DISULFID 32 56 BY SIMILARITY.
 FT DISULFID 33 72 BY SIMILARITY.
 FT CONFLICT 87 87 M -> L (IN REF. 1).
 SQ SEQUENCE 90 AA: 9987 MW: 50AF9679A267408F CRC64;
 Query Match 10.6%; Score 71.5; DB 1; Length 90;
 Best Local Similarity 27.1%; Pred. No. 0.32;
 Matches 23; Conservative 17; Mismatches 32; Indels 13; Gaps 5;
 Oy 6 LAIVLAVCAALHASEALLPIAS---SCTEVSHHSIRRLERVMKRIQADSDCLA 61
 Db 8 LAVLIAICVQTSAA---PVGSDPTSCCFT---YISRLQFPF-VADYYETNSQCPHA 59
 Oy 62 AVILHVKRRX-ICVSPHNHTVKOM 85
 Db 60 GVFTTRKRGVCANPDMVQDYIM 84
 RESULT 12
 CBP1_CAEEL STANDARD: PRT: 2056 AA.
 AC P34545;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein cbp-1.
 GN CBP-1 OR R10E11.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.

NCBI_TaxID=6239;
 OX 11 SEQUENCE FROM N.A.
 RP STRAIN-Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken A., Roopra A., Saunders D., Showkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Therry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RP REVISIONS, AND ALTERNATIVE SPLICING.
 RA Durbin R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a and b (shown here); may be
 CC produced by alternative splicing.
 CC -1- SIMILARITY: CONTAINS 1 ZN-TYPE ZINC FINGER.
 CC -1- SIMILARITY: CONTAINS 1 ZN-TYPE ZINC FINGER.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z29095; CAA82353.2; -
 DR EMBL: Z29095; CAA82353.2; -
 DR PIR: S40713; S40713.
 DR WormPep: R10E11.1a; CE28069.
 DR WormPep: R10E11.1b; CE21117.
 DR InterPro: IPR001487; Bromodomain.
 DR InterPro: IPR003101; KIX.
 DR InterPro: IPR001230; Ptenyl_site.
 DR InterPro: IPR000197; TAZ_finger.
 DR InterPro: IPR000433; Znf_ZZ.
 DR Pfam: PF00439; bromodomain; 1.
 DR Pfam: PF00569; Zz; 1.
 DR Pfam: PF02135; zf-TAZ; 2.
 DR Pfam: PF02172; KIX; 1.
 DR PRINTS: PR00503; BROMODOMAIN.
 DR SMART: SM00297; BROMO; 1.
 DR SMART: SM00291; Znf_ZZ; 1.
 DR PROSITE: PS00633; BROMODOMAIN_1; 1.
 DR PROSITE: PS00014; BROMODOMAIN_2; 1.
 DR PROSITE: PS01359; ZF_PHD_1; 1.
 DR PROSITE: PS01357; ZF_ZZ_1; 1.
 DR PROSITE: PS01355; ZF_ZZ_2; 1.
 DR Bromodomain; Metal-binding; Zinc; Zinc-finger; Alternative splicing.
 KW DOMAIN 881 953 BROMODOMAIN.
 FT ZN_FING 1493 1534 ZN-TYPE.
 FT DOMAIN 1687 2008 GLY/GIN-RICH.
 FT VARSPLIC 467 478 SPTQVTKKCSV -> F (IN ISOFORM A).
 SQ SEQUENCE 2056 AA; 227179 MW; 949FP4608C634F01 CRC64;
 Query Match 10.0%; Score 70; DB 1; Length 2056;
 Best Local Similarity 19.0%; Pred. No. 12;
 Matches 23; Conservative 26; Mismatches 60; Indels 12; Gaps 4;
 QY 13 VCAALHASEALIPASSCCTEVSHHSRLLEVRNMCRIQADGDCD-----LAAYLVH 67
 Db 1561 IASLVHACQ-----RDNCRHMSCHKMKRVYQTKMKC-KRINGTGPVCKQLALACQYH 1615

QY 68 KR-XRIVSPHNHTVKOMKVOAAXKNGKNCRRKHGKNSRAHOGKHETGAKT 125
 Db 1616 KCHTRDACTVPCFMNIRKILAEQKRSQORADMMRRMRMEGLQSHVGAAPSTVSNGT 1675
 QY 126 P 126
 Db 1676 P 1676
 RESULT 13
 Y19A_ECOLI STANDARD; PRT; 173 AA.
 ID Y19A_ECOLI
 AC Q47309;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Insertion element IS1397 hypothetical 20.1 kDa protein (ORF).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EPEC 25;
 RA Bachelier S., Clement J.M., Hofnung M., Gilson E.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE IS150/IS1296 ORFA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92970; CAA63546.1; -
 DR KM Transposable element; Hypothetical protein.
 SQ SEQUENCE 173 AA; 20098 MW; 82E9318A2E233A7C CRC64;
 Query Match 10.0%; Score 68; DB 1; Length 173;
 Best Local Similarity 27.4%; Pred. No. 1.5;
 Matches 29; Conservative 11; Mismatches 48; Indels 18; Gaps 4;
 QY 32 TEVSHHSRLLE-RVNMCRIOADGDCDLAAYLVH-----KRXICVSPHNHT 80
 Db 34 TSLSHWNLFLHGRALDCHRSYSPEDKLCVLYALGHSLSLPRVARFNIPISHN-T 92
 QY 81 VKOMKVOAAXKNGKNCRRKHGKNSRAHOGKHETGAKT 126
 Db 93 VKNMKIGYRKSGNEAFTRCKEK-----SMTRSDPTHEANMTP 132
 RESULT 14
 DSC2_HUMAN STANDARD; PRT; 901 AA.
 ID DSC2_HUMAN
 AC Q02487;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Desmocollin 2A/2B precursor (Desmosomal glycoprotein II and III)
 DE (Desmocollin-3).
 GN DSC2 OR DSC3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Keratinocytes;
 RX MEDLINE=91244819; PubMed=2037591;
 RA Parker A.E., Wheeler G.N., Arnemann J., Pidsley S.C., Ataliotis P.,
 RA Thomas C.L., Rees D.A., Magee A.I., Buxton R.S.;

RT "Desmosomal glycoproteins II and III. Cadherin-like junctional
 RL molecules generated by alternative splicing.";
 CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOSOME JUNCTIONS. INVOLVED
 CC IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE FILAMENTS
 CC MEDIATING CELL-CELL ADHESION. MAY CONTRIBUTE TO EPIDERMAL CELL
 CC POSITIONING (SPREAD/RETENTION) BY MEDIATING DIFFERENTIAL
 CC ADHESIVENESS BETWEEN CELLS THAT EXPRESS DIFFERENT ISOFORMS.
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 2A/DGII (SHOWN HERE) AND
 CC 2B/DGIII; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIA, MYOCARDIUM AND LYMPH
 CC NODES.
 CC -1- DOMAIN: CALCITONIN MAY BE BOUND BY THE CADHERIN-LIKE REPEATS
 CC (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56807; CAA40141.1; ALT_INIT.
 DR EMBL; X56807; CAA40142.1; ALT_INIT.
 DR PIR; A40390; IJHUBA.
 DR PIR; A40390; IJHUBA.
 DR HSP; P15116; INCJ.
 DR Genew; HGNC:3036; DSC2.
 DR MIM; 125645; .
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 4.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS50268; CADHERIN_2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
 KW Alternative splicing; Cytoskeleton; Calcium-binding; Phosphorylation.
 FT SIGNAL 1 27
 FT PROPEP 28 135
 FT CHAIN 136 901
 FT DOMAIN 136 901
 FT TRANSMEM 695 715
 FT DOMAIN 716 901
 FT DOMAIN 136 243
 FT DOMAIN 244 355
 FT DOMAIN 356 471
 FT DOMAIN 472 579
 FT DOMAIN 580 694
 FT CARBOHYD 34 34
 FT CARBOHYD 166 166
 FT CARBOHYD 392 392
 FT CARBOHYD 546 546
 FT CARBOHYD 629 629
 FT MOD_RES 864 864
 FT VARSPIC 837 847
 FT VARSPIC 847 864
 FT SEQUENCE 901 AA; 99961 MW; 30F7E3D33ECA67CC CRC64;
 SQ
 Query Match 9.9%; Score 67; DB 1; Length 901;
 Best Local Similarity 31.7%; Pred. No. 11;
 Matches 19; Conservative 11; Mismatches 18; Indels 12; Gaps 2;
 Oy 12 AYCAALAHSEALPIASCCTEVSHHT-----SRRLERVMNCRIQRADGCDLAAYTIIH 66
 Db 13 ALCRILLTLTALIFASADCKKNVTLHVPKSLDAEKLGRVNL-----KRCFTANLIH 65
 RESULT 15
 ST05_HUMAN

ID ST05_HUMAN STANDARD: PRT; 91 AA.
 AC P13501; Q43646; Q9N7A2;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Small inducible cytokine A5 precursor (CCIL5) (T-cell specific RANTES
 DE protein) (Sis-delta) (T cell-specific protein P228) (TCP228).
 GN SCYA5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88285659; PubMed-2456327;
 RA Schall T.J., Jongstra J., Dyer B.J., Jorgensen J., Clayberger C.,
 RA Davis M.W., Krensky A.M.;
 RT "A human T cell-specific molecule is a member of a new gene family.";
 RL J. Immunol. 141:1018-1025(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX Jang J.S., Kim B.E.;
 RA Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99228475; PubMed-10213461;
 RA Nomiyama H., Fukuda S., Ito M., Tanase S., Miura R., Yoshie O.;
 RT "Organization of the chemokine gene cluster on human chromosome
 RT 17q11.2 containing the genes for CC chemokine MIP-1, HCC-2, LRC, and
 RT RANTES.";
 RL J. Interferon Cytokine Res. 19:227-234(1999).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Zeng O.P., Yang R.Y., Fu L.C.;
 RA "The complete sequence of human beta-chemokine RANTES mRNA.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX TISSUE-Brain;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 49-56; 71-79 AND 83-91, AND FUNCTION.
 RX MEDLINE-96106406; PubMed-8525373;
 RA Cocchi F., DeVico A.L., Garzino-Demo A., Arya S.K., Gallo R.C.,
 RA Lusso P.;
 RT "Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major
 RT HIV-suppressive factors produced by CD8+ T cells.";
 RL Science 270:1811-1815(1995).
 RN [7]
 RP STRUCTURE BY NMR.
 RX MEDLINE-95352612; PubMed-7542919;
 RA Chung C.-W., Cooke R.M., Proudfoot A.E.I., Wells T.N.C.;
 RT "The three-dimensional solution structure of RANTES.";
 RL Biochemistry 34:9307-9314(1995).
 RN [8]
 RP STRUCTURE BY NMR.
 RX MEDLINE-95244456; PubMed-7537088;
 RA Skelton N.J., Aspitas F., Oger J., Schall T.J.;
 RT "Proton NMR assignments and solution conformation of RANTES, a
 RT chemokine of the C-C type.";
 RL Biochemistry 34:5329-5342(1995).
 RN [9]
 RP SYNTHESIS AND X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE-99111238; PubMed-9889151;
 RA Wilken J., Hoover D., Thompson D.A., Barlow P.N., McSparron H.,
 RA Plazat L., Wlodawer A., Lubkowski J., Kent S.B.;
 RT "Total chemical synthesis and high-resolution crystal structure of
 RT the potent anti-HIV protein AOP-RANTES.";
 RL Chem. Biol. 6:43-51(1999).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RA Hoover D.M., Shaw J., Gryczynski Z., Proudfoot A.E.I., Wells T.N.C.,

```
RA Lubkowski J.;
RT "The crystal structure of Met-RANTES: comparison with native RANTES
RL and AOP-RANTES."
CC Protein Pept. Lett. 7:73-82(2000).
CC -1- FUNCTION: CHEMOTACTANT FOR BLOOD MONOCYTES, MEMORY T HELPER
CC CELLS AND EOSINOPHILS. CAUSES THE RELEASE OF HISTAMINE FROM
CC BASOPHILS AND ACTIVATES EOSINOPHILS. BINDS TO CCRI, CCR3, CCR4 AND
CC CCR5. ONE OF THE MAJOR HIV-SUPPRESSIVE FACTORS PRODUCED BY CD8+ T
CC CELLS. RECOMBINANT RANTES PROTEIN INDUCES A DOSE-DEPENDENT
CC INHIBITION OF DIFFERENT STRAINS OF HIV-1, HIV-2, AND SIMIAN
CC IMMUNODEFICIENCY VIRUS (SIV).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: T-CELL AND MACROPHAGE SPECIFIC.
CC -1- INDUCTION: BY MITOGENS.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
CC C-C) (CHEMOKINE CC).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21121; AAA36725.1; -
CC EMBL; AF043341; AAC03541.1; -
CC EMBL; AF088219; AAC63331.1; -
CC EMBL; AF266753; AAF73070.1; -
CC EMBL; BC008600; AAH08600.1; -
CC PIR; A28815; A28815; -
CC PDB; 1HRJ; 14-OCT-96.
CC PDB; 1RTN; 03-JUN-95.
CC PDB; 1RTO; 03-JUN-95.
CC PDB; 1B3A; 23-APR-99.
CC PDB; 1EOT; 19-APR-00.
CC GeneW; HGNC:10632; SCYAS.
CC MIM; 187011; -
CC InterPro; IPR000827; CC_chemkine_sml.
CC InterPro; IPR001811; Chemokine_IL8.
CC Pfam; PF00048; IL8; 1.
CC SMART; SM00199; SCY; 1.
CC PROSITE; PS00472; SMALL_CYTOKINES_CC; 1.
CC Cytokine; Chemotaxis; T-cell; Signal; Inflammatory response;
CC 3D-structure.
CC KW SIGNAL 1 23
CC FT CHAIN 24 91 SMALL INDUCIBLE CYTOKINE A5.
CC FT DISULFID 33 57
CC FT DISULFID 34 73
CC FT CONFLICT 7 7 A -> R (IN REF. 1 AND 4).
CC FT CONFLICT 14 14 A -> V (IN REF. 4).
CC FT CONFLICT 14 14 F30BFAF9A87C620F CRC64;
CC SQ SEQUENCE 91 AA; 9990 MW; F30BFAF9A87C620F CRC64;

Query Match 9.7%; Score 66; DB 1; Length 91;
Best Local Similarity 27.7%; Pred. No. 1.3;
Matches 23; Conservative 16; Mismatches 28; Indels 16; Gaps 6;

QY 8 IVALAVCALHASEALPIAS--CCTEVSHIISRLERVMCRIGRADGCDLAAYI 64
DB 12 LIAFALCAPASAS----PYSDDTPCCFA---YIARP-LPRAHIKEFYTGKCSNPAVY 63
QY 65 LHVKKRXR-ICVSPHNHTVKOMK 86
DB 64 FVTRKNROYCANPE---KKWVR 82
```

Search completed: January 14, 2003, 18:13:47
Job time : 19.3444 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:15:35 ; Search time 16.9333 Seconds
(Without alignments)
145.508 Million cell updates/sec

Title: US-09-834-794A-1

Perfect score: 677

Sequence: 1 MOORGIAIVAVCAALHAS.....RNSNRAHQKHETGYHKTPY 127

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	673	99.4	127	9	US-09-898-751A-6
2	673	99.4	127	10	US-09-813-492-2
3	673	99.4	127	10	US-09-834-795A-1
4	673	99.4	127	10	US-09-931-381A-2
5	565	83.5	104	10	US-09-834-795A-2
6	418.5	61.8	130	9	US-09-898-751A-10
7	162.5	24.0	120	9	US-09-898-751A-14
8	162.5	24.0	120	9	US-09-792-793A-87
9	162.5	24.0	120	9	US-09-931-381A-12
10	158	23.3	164	9	US-09-898-751A-8
11	147.5	21.8	112	9	US-09-898-751A-12
12	147.5	21.8	112	10	US-09-831-381A-11
13	112	16.5	19	10	US-09-834-795A-5
14	98.5	14.5	149	10	US-09-272-162-2
15	97.5	14.4	149	12	US-10-000-759A-11
16	97	14.3	16	10	US-09-834-795A-4
17	95	14.0	150	12	US-10-000-759A-9
18	94	13.9	18	10	US-09-834-795A-3
19	94	13.9	150	10	US-09-834-795A-20

20	94	13.9	150	10	US-09-931-381A-13	Sequence 13, Appl
21	92	13.6	144	10	US-09-931-381A-14	Sequence 14, Appl
22	92	13.6	150	10	US-09-799-777-25	Sequence 25, Appl
23	70.5	10.4	651	12	US-10-153-064-133	Sequence 133, App
24	70.5	10.4	652	12	US-10-153-064-96	Sequence 96, Appl
25	70.5	10.4	652	12	US-10-153-064-99	Sequence 99, Appl
26	70.5	10.4	652	12	US-10-153-064-105	Sequence 105, App
27	70.5	10.4	652	12	US-10-153-064-132	Sequence 132, App
28	70.5	10.4	653	12	US-10-153-064-131	Sequence 131, App
29	70.5	10.4	656	12	US-10-153-064-130	Sequence 130, App
30	70.5	10.4	660	12	US-10-153-064-90	Sequence 90, Appl
31	70.5	10.4	660	12	US-10-153-064-93	Sequence 93, Appl
32	70.5	10.4	676	12	US-10-153-064-95	Sequence 95, Appl
33	70.5	10.4	676	12	US-10-153-064-98	Sequence 98, Appl
34	70.5	10.4	676	12	US-10-153-064-104	Sequence 104, App
35	70.5	10.4	676	12	US-10-153-064-127	Sequence 127, App
36	70.5	10.4	676	12	US-10-153-064-129	Sequence 129, App
37	70.5	10.4	677	12	US-10-153-064-125	Sequence 125, App
38	70.5	10.4	680	12	US-10-153-064-123	Sequence 123, App
39	70.5	10.4	684	12	US-10-153-064-92	Sequence 92, Appl
40	70.5	10.4	1184	12	US-10-153-064-89	Sequence 89, Appl
41	66	9.7	91	8	US-08-927-939-21	Sequence 21, Appl
42	66	9.7	91	9	US-10-057-275-8	Sequence 8, Appl
43	66	9.7	91	10	US-09-144-838-9	Sequence 9, Appl
44	66	9.7	91	10	US-09-834-795A-29	Sequence 29, Appl
45	66	9.7	91	12	US-10-158-366-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-898-751A-6

Sequence 6, Application US/09898751A

Patent No. US20020160024A1

GENERAL INFORMATION:

APPLICANT: Oldham, Elizabeth R.

APPLICANT: Soto, Hortensia

APPLICANT: Liu, Ying

APPLICANT: Hudak, Susan A.

APPLICANT: Morales, Janine M.

APPLICANT: Kellerman, Strid-Almee

APPLICANT: McEvoy, Leslie M.

APPLICANT: Bowman, Edward P.

APPLICANT: Zlotnik, Albert

TITLE OR INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS

FILE REFERENCE: DX0882X

CURRENT APPLICATION NUMBER: US/09/898, 751A

CURRENT FILING DATE: 2001-07-02

PRIOR APPLICATION NUMBER: US09/471,549

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: US60/136,570

PRIOR FILING DATE: 1999-05-27

PRIOR APPLICATION NUMBER: US60/113,858

PRIOR FILING DATE: 1998-12-24

NUMBER OF SEQ ID NOS: 16

SOFTWARE: PatentIn version 3.1

SEQ ID NO 6

LENGTH: 127

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (529)..(529)

OTHER INFORMATION: unknown amino; may be "A", "C", or "G"

US-09-898-751A-6

Query Match 99.4%: Score 673; DB 9; Length 127;

Best Local Similarity 98.4%: Pred. No. 1e-66;

Matches 125; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MOORGIAIVAVCAALHASAAILPIASSCCTEVSHHSIRLLERVMNCRIQADGCDL 60

```
|||||
Db 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Qy 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Db 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Qy 121 YGHKTPY 127
Db 121 YGHKTPY 127

RESULT 2
US-09-813-492-2
; Sequence 2, Application US/09813492
; Patent No. US2002009735A1
; GENERAL INFORMATION:
; APPLICANT: Labow, Mark A.
; APPLICANT: Bhalla, Umesh
; APPLICANT: Mickanin, Craig Stephen
; TITLE OF INVENTION: MAMMARY GLAND CHEMOKINE
; FILE REFERENCE: 12345
; CURRENT APPLICATION NUMBER: US/09/813,492
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 127
; TYPE: PRT
; ORGANISM: HUMAN
US-09-813-492-2

Query Match          99.4%; Score 673; DB 10; Length 127;
Best Local Similarity 98.4%; Pred. No. 1e-66;
Matches 125; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Db 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Qy 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Db 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Qy 121 YGHKTPY 127
Db 121 YGHKTPY 127

RESULT 3
US-09-834-795A-1
; Sequence 1, Application US/09834795A
; Patent No. US20020076710A1
; GENERAL INFORMATION:
; APPLICANT: Lawrence, Papsidero
; APPLICANT: Lyn, Dyster
; APPLICANT: Jana, Frustaci
; TITLE OF INVENTION: Detection and Treatment of Breast Cancer
; FILE REFERENCE: 3380/11127-US3
; CURRENT APPLICATION NUMBER: US/09/834,795A
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 09/146,580
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: 60/071,899
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 60/092,155
; PRIOR FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: (70)..(70)
;; OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
;; NAME/KEY: UNSURE
;; LOCATION: (91)..(91)
;; OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-834-795A-1

Query Match          99.4%; Score 673; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1e-66;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Db 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Qy 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Db 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Qy 121 YGHKTPY 127
Db 121 YGHKTPY 127

RESULT 4
US-09-931-381A-2
; Sequence 2, Application US/09931381A
; Patent No. US20020137107A1
; GENERAL INFORMATION:
; APPLICANT: Butcher, Eugene C.
; APPLICANT: Kunkel, Eric J.
; APPLICANT: Pan, Junliang
; APPLICANT: Soler-Ferran, Dulce
; TITLE OF INVENTION: Method for Identifying Agents which
; TITLE OF INVENTION: Modulate Chemokine "MCC"-Induced Functions of CCR3 and/or
; FILE REFERENCE: 1855.2010-003
; CURRENT APPLICATION NUMBER: US/09/931,381A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: U.S. 09/638,914
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(22)
US-09-931-381A-2

Query Match          99.4%; Score 673; DB 10; Length 127;
Best Local Similarity 98.4%; Pred. No. 1e-66;
Matches 125; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Db 1 MOORGLAIYALVCAALHASEALPIASSCTEVSHHISRLLERVMCRIOADGDCDL 60
Qy 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Db 61 AAVILHVKKRRICVSPHNHTVQMKVQAAXKNGKNCVCHRRKHHGKRSNRAHOGKHET 120
Qy 121 YGHKTPY 127
Db 121 YGHKTPY 127

RESULT 5
US-09-834-795A-2
```



```
; Sequence 2, Application US/09834795A
; Patent No. US20020076710A1
; GENERAL INFORMATION:
; APPLICANT: Lawrence, Papsidero
; APPLICANT: Lyn, Dyster
; APPLICANT: Jana, Frustack
; TITLE OF INVENTION: Detection and Treatment of Breast Cancer
; FILE REFERENCE: 3380/1127-US3
; CURRENT APPLICATION NUMBER: US/09/834,795A
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 09/146,580
; PRIOR FILING DATE: 1998-09-03
; PRIOR APPLICATION NUMBER: 60/071,899
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 60/092,155
; PRIOR FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
; NAME/KEY: UNSURE
; LOCATION: (68)..(68)
; OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
; US-09-834-795A-2
```

```
Query Match      83.5%; Score 565; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 5,2e-55;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 24 LPIASCCTEVSHHISRLRLERVMCRIOADGCDLAAYILHVKKRRCVSPHNHTVKQ 83
DB 1 LPIASCCTEVSHHISRLRLERVMCRIOADGCDLAAYILHVKKRRCVSPHNHTVKQ 60
QY 84 MMKVQAAKKNGKGNVCHRRKHHGRNSNRHAGKHETTYGHTPY 127
DB 61 MMKVQAAKKNGKGNVCHRRKHHGRNSNRHAGKHETTYGHTPY 104
```

RESULT 6

```
US-09-898-751A-10
; Sequence 10, Application US/09898751A
; Patent No. US20020160024A1
; GENERAL INFORMATION:
; APPLICANT: Oldham, Elizabeth R.
; APPLICANT: Soto, Hortensia
; APPLICANT: Liu, Ying
; APPLICANT: Hudak, Susan A.
; APPLICANT: Homey, Bernhard
; APPLICANT: Morales, Janine M.
; APPLICANT: Kellerman, Strid-Almee
; APPLICANT: Mcevoy, Leslie M.
; APPLICANT: Bowman, Edward P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
; FILE REFERENCE: DX0882XK
; CURRENT APPLICATION NUMBER: US/09/898,751A
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: US09/471,549
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US60/136,570
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US60/113,858
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; LENGTH: 130
```

```
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-898-751A-10
```

```
Query Match      61.8%; Score 418.5; DB 9; Length 130;
Best Local Similarity 63.2%; Pred. No. 6,7e-39;
Matches 79; Conservative 14; Mismatches 27; Indels 5; Gaps 1;
```

```
QY 1 MQGGLAIVLAVCAALHASEAILPIASCCTEVSHHISRLRLERVMCRIOADGCDL 60
DB 1 MQGGLTMAVAVCAVQTSFALIPMASCCTEVSHHISRLRLERVMCRIOADGCDL 60
QY 61 AAVILHVKKRRCVSPHNHTVKQMMKVQAAKKNGKGNVCHRRKHHGRNSNRHAGKHET 120
DB 61 AAVILHVKKRRCVSPHNHTVKQMMKVQAAKKNGKGNVCHRRKHHGRNSNRHAGKHET 115
QY 121 YGHTT 125
DB 116 RKHRT 120
```

RESULT 7

```
US-09-898-751A-14
; Sequence 14, Application US/09898751A
; Patent No. US20020160024A1
; GENERAL INFORMATION:
; APPLICANT: Oldham, Elizabeth R.
; APPLICANT: Soto, Hortensia
; APPLICANT: Liu, Ying
; APPLICANT: Hudak, Susan A.
; APPLICANT: Homey, Bernhard
; APPLICANT: Morales, Janine M.
; APPLICANT: Kellerman, Strid-Almee
; APPLICANT: Mcevoy, Leslie M.
; APPLICANT: Bowman, Edward P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
; FILE REFERENCE: DX0882XK
; CURRENT APPLICATION NUMBER: US/09/898,751A
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: US09/471,549
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US60/136,570
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US60/113,858
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 14
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-898-751A-14
```

```
Query Match      24.0%; Score 162.5; DB 9; Length 120;
Best Local Similarity 41.8%; Pred. No. 5,6e-11;
Matches 33; Conservative 14; Mismatches 29; Indels 3; Gaps 2;
```

```
QY 19 ASEAILPIAS--SCCTEVSHH-ISRLRLERVMCRIOADGCDLAAYILHVKKRRCV 75
DB 21 APEALPLPSSSTSCCTQLYRQPLPSRLRLRIVHMEIQADGCDLQAAVILHARRSCVH 80
QY 76 PHNHTVKQMMKVQAAKKNG 94
DB 81 PGNRLARWLERGGRKRLQG 99
```

RESULT 8

```
US-09-792-793A-87
; Sequence 87, Application US/09792793A
; Patent No. US20020168370A1
; GENERAL INFORMATION:
; APPLICANT: McDonald, John R.
```

```

      APPLICANT: Coggin, Phil
      TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING SECONDARY TISSUE DAM
      TITLE OF INVENTION: OTHER INFLAMMATORY CONDITIONS AND DISORDERS
      FILE REFERENCE: 25020-601D
      CURRENT APPLICATION NUMBER: US/09/792,793A
      CURRENT FILING DATE: 2001-02-22
      NUMBER OF SEQ ID NOS: 93
      SOFTWARE: PatentIn Ver. 2.0
      SEQ ID NO 87
      LENGTH: 120
      TYPE: PRT
      ORGANISM: Mus musculus
      FEATURE:
      OTHER INFORMATION: Mouse chemokine ALP
US-09-792-793A-87

Query Match          24.0%; Score 162.5; DB 9; Length 120;
Best Local Similarity 41.8%; Pred. No. 5,6e-11;
Matches 33; Conservative 14; Mismatches 29; Indels 3; Gaps 2;

OY 19 ASEAALPIAS--SCCTEVSHN-ISRRLERVMNCRIQADGDDLAIVLHVKKRXCIVS 75
      ||| ||| : ||||| : ||| : ||||| ||||| : |||
Db 21 APEALPLPSSSTSCQTQLRQPLPSRLRLRIYVMEIQEADGDCHLQAVVHLHARSVCVH 80
      | | : : ||| : |

OY 76 PHNHTVKQMKVQAAXKNG 94
      | | : : ||| : |
Db 81 PGNRSLARLERQGRKRLQG 99

RESULT 10
US-09-898-751A-8
; Sequence 8, Application US/09898751A
; Patent No. US2002016002A1
GENERAL INFORMATION:
; APPLICANT: Oldham, Elizabeth R.

```

```

      APPLICANT: Socio, Hortensia
      APPLICANT: Liu, Ying
      APPLICANT: Hudak, Susan A.
      APPLICANT: Homey, Bernhard
      APPLICANT: Morales, Janine M.
      APPLICANT: Kellerman, Sirlid-Almee
      APPLICANT: McEvoy, Leslie M.
      APPLICANT: Bowman, Edward P.
      APPLICANT: Zlotnik, Albert
      TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
      FILE REFERENCE: DX0882XK
      CURRENT APPLICATION NUMBER: US/09/898,751A
      CURRENT FILING DATE: 2001-07-02
      PRIOR APPLICATION NUMBER: US09/471,549
      PRIOR FILING DATE: 1999-12-23
      PRIOR APPLICATION NUMBER: US60/136,570
      PRIOR FILING DATE: 1999-05-27
      PRIOR APPLICATION NUMBER: US60/113,858
      PRIOR FILING DATE: 1998-12-24
      NUMBER OF SEQ ID NOS: 16
      SOFTWARE: PatentIn version 3.1
      SEQ ID NO 8
      LENGTH: 164
      TYPE: PRF
      ORGANISM: Homo sapiens
US-09-898-751A-8

Query Match          23.3%; Score 158; DB 9; Length 164;
Best Local Similarity 40.8%; Pred. No. 2,5e-10;
Matches    31; Conservative   16; Mismatches   27; Indels     2; Gaps         2

OY       21 EAI-LPIASCCTEVSHH-ISRRLLEVMNCRIQADGDCDLAAVLILHYKKRYICVSPHN 78
        |||::||::||||::: : || | :: : || || || ||:||: |: || || ||
Db      68 ELALPFSSTSCCTQLYROPPLSLRLLRIVHMELQPADGDCHLAQAVLVHLARSVCVHPQN 127

OY       79 HTVKOMMKVVOAXKNKG 94
        :::::|
Db      128 RSLANMLERGGCKRIQG 143

--RESULT 11
US-09-898-751A-12
Sequence 12, Application US/09898751A
Patent No. US20020160024A1
GENERAL INFORMATION:
APPLICANT: Oldham, Elizabeth R.
APPLICANT: Soco, Hortensia
APPLICANT: Liu, Ying
APPLICANT: Hudak, Susan A.
APPLICANT: Homey, Bernhard
APPLICANT: Morales, Janine M.
APPLICANT: Kellerman, Sirlid-Almee
APPLICANT: McEvoy, Leslie M.
APPLICANT: Bowman, Edward P.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
FILE REFERENCE: DX0882XK
CURRENT APPLICATION NUMBER: US/09/898,751A
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: US09/471,549
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: US60/136,570
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US60/113,858
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 112
TYPE: PRF
ORGANISM: Homo sapIens
US-09-898-751A-12
```



```
Sequence 11, Application US/10000759A
Patent No. US2002014191A1
GENERAL INFORMATION:
APPLICANT: Andrew, David P.
APPLICANT: Zabel, Brian A.
APPLICANT: Poonath, Paul D.
TITLE OF INVENTION: ANTI-GPR-9-6 ANTIBODIES AND METHODS OF
IDENTIFYING MODULATORS OF GPR-9-6 FUNCTION
FILE REFERENCE: 1855.1064-003
CURRENT APPLICATION NUMBER: US/10/000,759A
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: US/09/522,752
PRIOR FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: US 09/266,464
PRIOR FILING DATE: 1999-03-11
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 11
LENGTH: 149
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (104)...(104)
OTHER INFORMATION: Xaa- Met or Thr
US-10-000-759A-11

Query Match 14.4%; Score 97.5; DB 12; Length 149;
Best Local Similarity 30.5%; Pred. No. 0.00091;
Matches 29; Conservative 14; Mismatches 41; Indels 11; Gaps 4;

OY 26 IASSCCTEVSHHSRLLEVRNMCRIQRADGCDLAVALHV-KRKR-ICVSPHNHTYKQ 83
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 VEDDCCLAVHYPIGWAVLRRAWTYRIQEVSGSCNLPAAIFYLPKRHRKVCNPKSREYQR 85
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

OY 84 WMKYQAAXKNGKGNVCHRRKKHGGKRNRAHOGKH 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 AMKLIDAR----NKVFALHH---NKOTFGPH 111
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: January 14, 2003, 18:27:27
Job time : 16.9333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:11:25 ; Search time 28.2222 Seconds
(without alignments)
132.403 Million cell updates/sec

Title: US-09-834-794a-1

Perfect score: 677
Sequence: 1 MOORGLAIVLAVCAALHAS.....RNSNRAHQKHETGYHETPY 127Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	99.4	137	4	US-09-146-580-1
2	565	83.5	104	4	US-09-146-580-2
3	112	16.5	19	4	US-09-146-580-5
4	98.5	14.5	149	2	US-08-874-460-2
5	97	14.3	16	4	US-09-146-580-4
6	94	13.9	18	4	US-09-146-580-3
7	74.5	11.0	96	3	US-09-071-353-2
8	74.5	11.0	121	5	PCT-US95-07171-3
9	70	10.3	71	3	US-09-071-353-3
10	67	9.9	96	4	US-09-230-637-44
11	66	9.7	90	1	US-09-230-637-40
12	66	9.7	91	1	US-08-347-492B-12
13	66	9.7	91	1	US-08-375-346A-5
14	66	9.7	91	1	US-08-480-449-21
15	66	9.7	91	1	US-08-653-682-3
16	66	9.7	91	2	US-08-421-144A-8
17	66	9.7	91	2	US-08-660-542-21
18	66	9.7	91	2	US-08-798-143-12
19	66	9.7	91	2	US-08-467-123B-5
20	66	9.7	91	3	US-08-936-772-3
21	66	9.7	91	4	US-08-836-922-14
22	66	9.7	91	4	US-09-395-918-3
23	66	9.7	91	4	US-08-679-493A-155
24	66	9.7	91	4	US-08-479-603-21
25	66	9.7	91	4	US-09-230-371A-25
26	65	9.6	110	1	US-08-359-696-2
27	65	9.6	120	1	US-08-347-492B-2

28	65	9.6	120	2	US-08-798-143-2	Sequence 2, Appl1
29	65	9.6	120	5	PCT-US95-15484-2	Sequence 2, Appl1
30	65	9.6	136	5	PCT-US95-07171-2	Sequence 2, Appl1
31	64.5	9.5	104	4	US-09-133-521-2	Sequence 2, Appl1
32	63.5	9.4	109	1	US-08-352-324A-2	Sequence 2, Appl1
33	63.5	9.4	109	1	US-08-352-324A-9	Sequence 2, Appl1
34	63.5	9.4	109	2	US-08-862-607-2	Sequence 2, Appl1
35	63.5	9.4	109	2	US-08-862-607-9	Sequence 2, Appl1
36	63.5	9.4	109	3	US-09-203-235-2	Sequence 2, Appl1
37	63.5	9.4	109	3	US-09-203-235-9	Sequence 2, Appl1
38	63.5	9.4	109	3	US-08-982-493-4	Sequence 2, Appl1
39	63.5	9.4	109	4	US-08-943-336A-4	Sequence 2, Appl1
40	63.5	9.4	109	5	PCT-US95-01780-4	Sequence 2, Appl1
41	63.5	9.4	109	5	PCT-US95-16144-2	Sequence 2, Appl1
42	63.5	9.4	109	5	PCT-US95-16144-9	Sequence 2, Appl1
43	63	9.3	86	2	US-08-421-144A-9	Sequence 2, Appl1
44	63	9.3	93	2	US-08-633-682-2	Sequence 2, Appl1
45	63	9.3	93	3	US-08-936-772-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-146-580-1
Sequence 1, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papsidero, Lawrence D
APPLICANT: Dyster, Lyn M
APPLICANT: Frustaci, Jana M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
CURRENT FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: UNSURE
LOCATION: (70)
OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
FEATURE:
NAME/KEY: UNSURE
LOCATION: (91)
OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-146-580-1
Query Match 99.4%: Score 673; DB 4; Length 127;
Best Local Similarity 100.0%: Pred. No. 1; le-76;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MOORGLAIVLAVCAALHASAIIPIASSCCTEYSHHSRLLERVNNCRIGADGCDL 60
Db 1 MOORGLAIVLAVCAALHASAIIPIASSCCTEYSHHSRLLERVNNCRIGADGCDL 60
QY 61 AAVILHVKRRIKICVSPHNHTVKOMKVOAAKNGKGVNCHRRKNSRAHQGHET 120
Db 61 AAVILHVKRRIKICVSPHNHTVKOMKVOAAKNGKGVNCHRRKNSRAHQGHET 120
QY 121 YGHKTPY 127
Db 121 YGHKTPY 127
RESULT 2
```

US-09-146-580-2
; Sequence 2, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; EARLIER FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)
; OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
US-09-146-580-2

Query Match 83.5%; Score 565; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 2.7e-63;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 LPASSCCTEVSHHSRLRLRYVMNCRIORADGCDLAIVLHVKKRXCSPHNHTVQ 83
DB 1 LPASSCCTEVSHHSRLRLRYVMNCRIORADGCDLAIVLHVKKRXCSPHNHTVQ 60

OY 84 WKKVQAAXKNGKGNVCHRRKHHGKRSNRAHQKHETYGHTPY 127
DB 61 WKKVQAAXKNGKGNVCHRRKHHGKRSNRAHQKHETYGHTPY 104

RESULT 3
US-09-146-580-5
; Sequence 5, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 19
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-146-580-5

Query Match 16.5%; Score 112; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 NSNRHQKHETYGHTPY 127

DB 1 NSNRHQKHETYGHTPY 19
|||||

RESULT 4
US-08-874-460-2
; Sequence 2, Application US/08874460
; Patent No. 5981231
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-FEI
; APPLICANT: KREIDER, BRENT
; APPLICANT: ROSEN, CRAIG
; TITLE OF INVENTION: CHEMOKINE BETA 15
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,460
; FILING DATE: HERewith
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/019,837
; FILING DATE: 17-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0420001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 149 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-874-460-2

Query Match 14.5%; Score 98.5; DB 2; Length 149;
Best Local Similarity 30.5%; Pred. No. 9.5e-05;
Matches 29; Conservative 14; Mismatches 41; Indels 11; Gaps 4;

OY 26 IASSCCTEVSHHSRLRLRYVMNCRIORADGCDLAIVLHV-KRR-ICVSPHNHTVQ 83
DB 26 VEDDCCLAVHYPLGAVLVRAMTYRIQEVSGSCNLPALFYLPKRHRKVCNPKSEVOR 85

OY 84 WKKVQAAXKNGKGNVCHRRKHHGKRSNRAHQKH 118
DB 86 AMKLIDAR-----NKVFALHH-----NQTFFGPH 111

RESULT 5
US-09-146-580-4
; Sequence 4, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03

EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-146-580-4

Query Match 14.3%; Score 97; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 92 KNGKGNVCHRRKHGK 107
DB 1 KNGKGNVCHRRKHGK 16

RESULT 6
US-09-146-580-3
Sequence 3, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papadiero, Lawrence D
APPLICANT: Dyster, Lyn M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
EARLIER FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
US-09-146-580-3

Query Match 13.9%; Score 94; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 TEVSHHSRLLERYVMC 49
DB 1 TEVSHHSRLLERYVMC 18

RESULT 7
US-09-071-353-2
Sequence 2, Application US/09071353
Patent No. 6057426
GENERAL INFORMATION:
APPLICANT: Lesslauer, Werner
APPLICANT: Utans-Schneitz, Urtike
TITLE OF INVENTION: NEW CHEMOKINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: N.J.
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 97107135.2
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kreisler, Lewis J
REGISTRATION NUMBER: 38522
REFERENCE/DOCKET NUMBER: 13235
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 235-4387
TELEFAX: (973) 235-2363
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-353-2

Query Match 11.0%; Score 74.5; DB 3; Length 96;
Best Local Similarity 31.5%; Pred. No. 0.055;
Matches 29; Conservative 13; Mismatches 35; Indels 15; Gaps 6;

OY 1 MOQRGLAVALA-VCAIAHASEALPIASCC---TEVSHHSRLLERYVMCRIGRADG 56
DB 1 MACKHLPEPLAAGVLAIVGCSOSEAASNFDCCLTYTKNVYHHARNFVGTFT---QMADE 56

OY 57 DCDLAIVILHYKRXR-ICVSPHNHTVKO-WMK 86
DB 57 ACIDNAILFIHLKSRKSCVADP-----KQIWK 83

RESULT 8
PCT-US95-07171-3
Sequence 3, Application PC/TUS9507171
GENERAL INFORMATION:
APPLICANT: MOONEY, J.
TITLE OF INVENTION: Human Chemokine Beta-12
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAYN, GILFILLAN,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07171
FILING DATE: 6 JUN 95
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: None
FILING DATE: None
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 191 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
PCT-US95-07171-3

Query Match 11.0%; Score 74.5; DB 5; Length 191;
Best Local Similarity 22.3%; Pred. No. 0.13;
Matches 31; Conservative 24; Mismatches 49; Indels 35; Gaps 6;

QY 8 IVALAVCALHASEALIP-----IASSCCTEVSHHI-SRRLLEVNMCRIORADGDDDLAA 62
DB 27 LVLLITTSASRSQPVPEKVNTPSTCCLAYEKLPRRLVVG-----YKALNCHLPA 80
QY 63 VILHVRXR-ICVSPHNHTVKOMKYO-----AAXKNGKGVCHRRKHNG 106
DB 81 IIVTRNRREVCTNPNDMDVQEKIKDNPRLPLTRNLSTYKITTANQOQLNSQXXPG 140
QY 107 -----KNSNRHAQOK 117
DB 141 FSGSPCLQKRGVNLXKQOK 159

RESULT 9
US-09-071-353-3
Sequence 3, Application US/09071353
Patent No. 6057426
GENERAL INFORMATION:
APPLICANT: Leaslauer, Werner
TITLE OF INVENTION: NEW CHEMOKINE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: N.J.
COUNTRY: U.S.A.
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 97107135.2
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Kreisler, Lewis J
REGISTRATION NUMBER: 38522
REFERENCE/DOCKET NUMBER: 13235
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 235-4387
TELEFAX: (973) 235-2363
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-353-3

Query Match 10.3%; Score 70; DB 3; Length 71;
Best Local Similarity 33.9%; Pred. No. 0.14;
Matches 21; Conservative 9; Mismatches 18; Indels 14; Gaps 5;
QY 30 CC-----TEVSHHISRLLERVMNCRIORADGDDDLAAVILHVRXR-ICVSPHNHTVKO-W 84

DB 6 CCLTYTKNYYHARNFVGFTT-----QMADEACDINAIFHLKSRSCVADP-----KOIW 56
QY 85 MK 86
DB 57 VK 58

RESULT 10
US-09-230-637-44
Sequence 44, Application US/09230637
Patent No. 6264958
GENERAL INFORMATION:
APPLICANT: Hayward, Gary
APPLICANT: Nicholas, John
APPLICANT: Hardwick, J. Marie
APPLICANT: Reitz, Marvin
TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
FILE REFERENCE: 1107.78372
CURRENT APPLICATION NUMBER: US/09/230,637
CURRENT FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: 60/022,591
PRIOR FILING DATE: 1996-07-25
PRIOR APPLICATION NUMBER: PCT US 97/12931
PRIOR FILING DATE: 1997-07-24
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 44
LENGTH: 96
TYPE: PRT
ORGANISM: Homo sapiens
US-09-230-637-44

Query Match 9.9%; Score 67; DB 4; Length 96;
Best Local Similarity 29.4%; Pred. No. 0.47;
Matches 25; Conservative 16; Mismatches 34; Indels 10; Gaps 5;

QY 7 AIVALVCAALHASEALIPAS---SCCTEVSHHISRRL-LEVRNMCRIORADGDDDLAA 62
DB 6 ALIMLLILAAFPDGLTGPASVPTCCFNLA---NRKIPLORLESYR--RISGKCPQKA 60
QY 63 VILHVRXR-ICVSPHNHTVKOMK 86
DB 61 VIFKRLARDICADPKRRWVDSMK 85

RESULT 11
US-09-230-637-40
Sequence 40, Application US/09230637
Patent No. 6264958
GENERAL INFORMATION:
APPLICANT: Hayward, Gary
APPLICANT: Nicholas, John
APPLICANT: Hardwick, J. Marie
APPLICANT: Reitz, Marvin
TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma
FILE REFERENCE: 1107.78372
CURRENT APPLICATION NUMBER: US/09/230,637
CURRENT FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: 60/022,591
PRIOR FILING DATE: 1996-07-25
PRIOR APPLICATION NUMBER: PCT US 97/12931
PRIOR FILING DATE: 1997-07-24
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40
LENGTH: 90
TYPE: PRT
ORGANISM: Homo sapiens
US-09-230-637-40


```

RESULT 14
US-08-480-449-21
; Sequence 21, Application US/08480445
; Patent No. 5688927
; GENERAL INFORMATION:

```

```

1 APPLICANT: Godiska, Ronald
2 APPLICANT: Gray, Patrick W.
3 TITLE OF INVENTION: MACROPHAGE DERIVED CHEMOKINE
4 NUMBER OF SEQUENCES: 24
5 CORRESPONDENCE ADDRESS:
6 ADDRESSSEE: Marshall, O'Toole, Gerstein, Murray & Borun
7 STREET: 6300 Sears Tower, 233 South Wacker Drive
8 CITY: Chicago
9 STATE: Illinois
10 COUNTRY: United States of America
11 ZIP: 60606-6402
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: Floppy disk
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: PC-DOS/MS-DOS
17 SOFTWARE: PatentIn Release #1.0, Version #1.25
18
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/480, 449
21 FILING DATE:
22 CLASSIFICATION: 530
23
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Gass, David A.
26 REGISTRATION NUMBER: 38,153
27 REFERENCE/DOCKET NUMBER: 27866/32779
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: 312/474-6300
30 TELEFAX: 312/474-0448
31
32 TELEX: 25-3856
33 INFORMATION FOR SEQ ID NO: 21:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 91 amino acids
36 TYPE: amino acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 MOLECULE TYPE: peptide
40
41 FEATURE:
42 NAME/KEY: misc_feature
43 OTHER INFORMATION: "RANTES"
44
45 US-08-480-449-21

```

	9.7%	Score	66;	DB	1;	Length	91;	
Query Match	Similarity	27.7%	Pred.	No.	0.59;			
Best Local	Conservative	16;	Mismatches	28;	Indels	16;	Gaps	6;
OY	8	I VALVCAALIHSEAILPIAS---	CCTEVSHHSIRLLERYMCRIGRADCDLAAYI	64				
		::: ::	::	:::	::	::		
Dd	12	LIALTLCAIPASAS----	PYSDDTPCCFA--YIANP-LPRHAIKEYFTSGKSNPAVV	63				
OY	65	LHVXR-X-ICVSPHNHTVKOMX	86					
		::: ::	::					
Dd	64	FVTRKNROYCANFE----	KKWVR	82				

RESULT 15
 US-08-633-682-3
 : Sequence 3, Application US/08633682
 : Patent No. 5840544
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Hawkins, Phillip R.
 : APPLICANT: Bandman, Olga
 : APPLICANT: Murry, Lynn E.
 : TITLE OF INVENTION: NOVEL RANTES HOMOLOG FROM PROSTATE
 :
 : NUMBER OF SEQUENCES: 5
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 :
 : COUNTRY: U.S.
 : ZIP: 94304
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible

```

? OPERATING SYSTEM: DOS
? SOFTWARE: FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/633, 682
? FILING DATE: Filed Herewith
? ATTORNEY/AGENT INFORMATION:
? NAME: Luther, Barbara J
? REGISTRATION NUMBER: 33,954
? REFERENCE/DOCKET NUMBER: PF-0063 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-855-0555
? TELEFAX: 415-852-0195
? INFORMATION FOR SEQ ID NO: 3:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 91 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? IMMEDIATE SOURCE:
? LIBRARY: Genbank
? CLONE: GI 134510
?
? US-08-633-682-3

```

[illegible]

Search completed: January 14, 2003, 18:17:03
Job time : 29.2222 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:07:25 : Search time 9.60556 Seconds
(without alignments)
407.566 Million cell updates/sec

Title: US-09-834-794A-5
Perfect score: 112
Sequence: 1 NSNRAHQGRHETYGKTPY 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	50	44.6	899	3	08TFN6	08TFN6 collettoria
2	48	42.9	567	5	096156	096156 plasmodium
3	47	42.0	735	5	09NES7	09NES7 caenorhabdi
4	46	41.1	770	16	054071	054071 talstonia s
5	46	41.1	1403	11	09QXG1	09QXG1 mus musculu
6	45	40.2	109	2	09K186	09K186 bacillus an
7	45	40.2	109	2	09K185	09K185 bacillus an
8	45	40.2	112	2	09K184	09K184 bacillus an
9	45	40.2	115	2	09K183	09K183 bacillus an
10	45	40.2	118	2	09K182	09K182 bacillus an
11	45	40.2	118	2	09K181	09K181 bacillus an
12	45	40.2	121	2	09K180	09K180 bacillus an
13	45	40.2	124	2	09K179	09K179 bacillus an
14	45	40.2	124	2	09K178	09K178 bacillus an
15	45	40.2	130	2	09K177	09K177 bacillus an
16	45	40.2	133	2	09K176	09K176 bacillus an

17	45	40.2	192	15	041793	041793 human immun
18	45	40.2	265	2	09K191	09K191 bacillus an
19	45	40.2	285	16	09FBJ7	09FBJ7 streptomyces
20	45	40.2	381	5	09XZV5	09XZV5 geodia cydo
21	45	40.2	382	5	019662	019662 caenorhabdi
22	45	40.2	427	10	09LX17	09LX17 arabidopsi
23	45	40.2	654	2	09F2M4	09F2M4 chromaliu
24	45	40.2	897	5	060961	060961 leishmania
25	45	40.2	1035	5	09V853	09V853 drosophila
26	45	40.2	1061	5	09U3W2	09U3W2 drosophila
27	44.5	39.7	289	2	046527	046527 bacteroides
28	44	39.3	165	9	038430	038430 bacterioph
29	44	39.3	193	15	0908K9	0908K9 human immun
30	44	39.3	509	5	094888	094888 drosophila
31	44	39.3	509	5	09VZM8	09VZM8 drosophila
32	44	39.3	522	5	09XTK9	09XTK9 drosophila
33	44	39.3	722	5	061208	061208 caenorhabdi
34	44	39.3	929	5	019300	019300 caenorhabdi
35	44	39.3	1033	5	044323	044323 forficula a
36	44	39.3	1130	4	09C0D6	09C0D6 homo sapien
37	43.5	38.8	533	5	09N755	09N755 plasmodium
38	43.5	38.8	597	5	09NC64	09NC64 plasmodium
39	43.5	38.8	654	5	09TY99	09TY99 plasmodium
40	43	38.4	155	17	028417	028417 archaeoglob
41	43	38.4	294	2	09K187	09K187 bacillus ce
42	43	38.4	296	2	09K189	09K189 bacillus ce
43	43	38.4	555	10	080469	080469 arabidopsi
44	43	38.4	784	5	09CPH1	09CPH1 calliphora
45	43	38.4	792	16	09ZJ13	09ZJ13 helicobacte

ALIGNMENTS

RESULT 1

Q8TFN6 ID Q8TFN6 PREDIMINARY; PRT; 899 AA.
AC 08TFN6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Chitin synthase A.
GN CHSA.
OS Colletotrichum graminicola (Anthracose fungus) (Glomerella graminicola).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes; Phyllachorales; Phyllachoraceae;
OC Glomerella.
OC NCBI_TaxID=31870;
RN [1]
RP SEQUENCE FROM N.A.
RA Werner S., Deising H.B.;
RT "Chitin synthases of the maize pathogen Colletotrichum graminicola";
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052545; MAL23717.1;
SQ SEQUENCE 899 AA; 101416 MW; A5E46C8E0EC70925 CRC64;

Query Match 44.6%; Score 50; DB 3; Length 899;
Best Local Similarity 42.9%; Pred. No. 9.7;
Matches 9; Conservative 6; Mismatches 4; Indels 2; Gaps 1;

QY 1 NSNRAHQ--GRHETYGKTPY 19
| | | | | : | | | | |
Db 27 NNNAAKPPPPQHDPRGHTSPH 47

RESULT 2

096156 ID 096156 PRELIMINARY; PRT; 567 AA.
AC 096156;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)

DE Membrane transporter.
GN PF80275W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Garner M.J., Tetteil H., Carucci D.J., Cummings L.M., Aravind L.,
Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Petrea M.,
Salberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
Frazer C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
falciparum.";
RL Science 282:1126-1132(1998).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AE001384; AAC71844.1; -;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 567 AA; 62906 MW; 210AFC5C8A7174E8 CRC64;

Query Match 42.9%; Score 48; DB 5; Length 567;
Best Local Similarity 52.9%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGYGH 17
DB 46 NSNRHOGKHETGYGH 62

RESULT 3
O9NES7 PRELIMINARY; PRT; 735 AA.
ID O9NES7;
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Y3986B_gg protein (Y3986A.pp protein).
GN Y3986B.GG OR Y3986A.pp.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Sulston J.E.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Sulston J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL133895; CAB60938.1; -;
DR EMBL: AL133948; CAC51077.1; -;
DR HSP; P13231; 1HCE.
DR InterPro: IPR002395; Kininogen.
DR PRINTS: PRO0334; KININOGEN.
SQ SEQUENCE 735 AA; 80255 MW; 6EF7B831DBE41159 CRC64;

Query Match 42.0%; Score 47; DB 5; Length 735;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 AHOGKHETGYGH 15
DB 11 11111111

DB 437 AHGHGHHSHGH 447

RESULT 4
ID 054071 PRELIMINARY; PRT; 770 AA.
AC 054071;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE PRNA protein precursor (Outer membrane receptor PRNA signal peptide
protein).
GN PRNA OR RSP0878 OR RS01649.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM1100;
RX MEDLINE=98143429; PubMed=9484898;
RA Marenza M.S., Brito B.M., Callard D., Genin S., Barberis P.A.,
Boucher C.A., Ariat M.;
RT "Prna controls a novel regulatory pathway required for the specific
induction of Ralstonia solanacearum hrp genes in the presence of plant
cells.";
RL Mol. Microbiol. 27:437-453(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-GM1100; PLASMID-MEGAPLASMID;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
Ariat M., Billault A., Brotier P., Camus J.C., Cattolico L.,
Chandler M., Choisme N., Claudel-Renard C., Cunac S., Demange N.,
Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,
Sigler P., Thebaud P., Whalen M., Wincker P., Levy M.,
Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AJ001084; CAA04519.1; -;
DR EMBL: AL646081; CAD18029.1; -;
KW Signal; Plasmid; Complete proteome.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 770 PRNA PROTEIN.
SQ SEQUENCE 770 AA; 82310 MW; B0D1EFAFSBADS5E CRC64;

Query Match 41.1%; Score 46; DB 16; Length 770;
Best Local Similarity 46.7%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGYGH 15
DB 304 NSDAGHTGRRTSYGH 318

RESULT 5
ID 090XG1 PRELIMINARY; PRT; 1403 AA.
AC 090XG1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MucN Muc4 (Fragment).
GN MUC4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartman A.E., Shekels L.L., Anway R.E., Gibson I.K., Mocca R.,
Ho S.B.;
RT "Identification and structure of a Mouse Homolog to the Human MUC4

RT Gene
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF218265; AAF23818.1; -
DR MGD: MGI:2153525; Mus4.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003886; Nidogen_ext.
DR InterPro: IPR001846; WVF_D.
DR Pfam: PF00094; vwd; 1.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00001; EGF-like; 2.
DR SMART: SM00539; NIDO; 1.
DR SMART: SM00216; VMD; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 1403 AA; 154669 MW; 0A51AFB7ED52EA7C CRC64;

OY 5 AHOGKHETYGKTP 18
Db 216 AHQSQHTYGTGP 229
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 6
O9K186 PRELIMINARY; PRT; 109 AA.
ID O9K186
AC O9K186
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2PT;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238889; AAF86204.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 109
SQ SEQUENCE 109 AA; 12873 MW; 3E92FF8D6C971E7A CRC64;

OY 2 SNRAHOGKHETYG 15
Db 63 THOGHGHHDHGH 76
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Query Match 40.2%; Score 45; DB 2; Length 109;
Best Local Similarity 42.9%; Pred. No. 6.2;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 7
O9K185 PRELIMINARY; PRT; 109 AA.
ID O9K185
AC O9K185
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA0015;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238890; AAF86205.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 109
SQ SEQUENCE 109 AA; 12979 MW; 42AFA1655C3BB273 CRC64;

OY 2 SNRAHOGKHETYG 15
Db 60 THOGHGHHDHGH 73
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Query Match 40.2%; Score 45; DB 2; Length 109;
Best Local Similarity 42.9%; Pred. No. 6.2;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 8
O9K184 PRELIMINARY; PRT; 112 AA.
ID O9K184
AC O9K184
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A46;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238891; AAF86206.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 112
SQ SEQUENCE 112 AA; 13293 MW; 25FE4876D3D7E193 CRC64;

OY 2 SNRAHOGKHETYG 15
Db 72 THOGHGHHDHGH 85
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Query Match 40.2%; Score 45; DB 2; Length 112;
Best Local Similarity 42.9%; Pred. No. 6.4;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 9
O9K183 PRELIMINARY; PRT; 115 AA.
ID O9K183
AC O9K183
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA0015;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238892; AAF86207.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 115
SQ SEQUENCE 115 AA; 13593 MW; 25FE4876D3D7E193 CRC64;

RP SEQUENCE FROM N.A.
RC STRAIN=MOZ-3;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238890; AAF86205.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 109
SQ SEQUENCE 109 AA; 12979 MW; 42AFA1655C3BB273 CRC64;

OY 2 SNRAHOGKHETYG 15
Db 60 THOGHGHHDHGH 73
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Query Match 40.2%; Score 45; DB 2; Length 109;
Best Local Similarity 42.9%; Pred. No. 6.2;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 8
O9K184 PRELIMINARY; PRT; 112 AA.
ID O9K184
AC O9K184
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A46;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238891; AAF86206.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 112
SQ SEQUENCE 112 AA; 13293 MW; 25FE4876D3D7E193 CRC64;

OY 2 SNRAHOGKHETYG 15
Db 72 THOGHGHHDHGH 85
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Query Match 40.2%; Score 45; DB 2; Length 112;
Best Local Similarity 42.9%; Pred. No. 6.4;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

RESULT 9
O9K183 PRELIMINARY; PRT; 115 AA.
ID O9K183
AC O9K183
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
ON NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA0015;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Kelm P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238892; AAF86207.1; -
DR HSSP: P13231; IHCE.
FT NON_TER 115
SQ SEQUENCE 115 AA; 13593 MW; 25FE4876D3D7E193 CRC64;

RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238892; AAF86207.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.
DR PRINTS: PR00334; KININOGEN.
FT NON_TER 115
SQ SEQUENCE 115 AA; 13616 MW; 55D5BA3807AEB183 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 115;
Best Local Similarity 42.9%; Pred. No. 6.6;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 75 THQGHGHHHHGH 88

RESULT 10
ID 09K182 PRELIMINARY; PRT; 118 AA.
AC 09K182;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VOLUM;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238893; AAF86208.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.
DR PRINTS: PR00334; KININOGEN.
FT NON_TER 118
SQ SEQUENCE 118 AA; 13850 MW; 17D2BF9F1F32B5A3 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 118;
Best Local Similarity 42.9%; Pred. No. 6.8;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 75 THQGHGHHHHGH 88

RESULT 11
ID 09K181 PRELIMINARY; PRT; 118 AA.
AC 09K181;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A24;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238894; AAF86209.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.
DR PRINTS: PR00334; KININOGEN.
FT NON_TER 118
SQ SEQUENCE 118 AA; 13947 MW; 1FF1609C54178C59 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 118;
Best Local Similarity 42.9%; Pred. No. 6.8;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 72 THQGHGHHHHGH 85

RESULT 12
ID 09K180 PRELIMINARY; PRT; 121 AA.
AC 09K180;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ZIM69;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238895; AAF86210.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.
DR PRINTS: PR00334; KININOGEN.
FT NON_TER 121
SQ SEQUENCE 121 AA; 14269 MW; 85F5E911130CFC72 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 121;
Best Local Similarity 42.9%; Pred. No. 7;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 75 THQGHGHHHHGH 88

RESULT 13
ID 09K179 PRELIMINARY; PRT; 124 AA.
AC 09K179;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BA1015;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238896; AAF86211.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.

DR EMBL: AF238894; AAF86209.1; -.
DR HSSP: P13231; IHCE.
DR NON_TER 118
SQ SEQUENCE 118 AA; 13947 MW; 1FF1609C54178C59 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 118;
Best Local Similarity 42.9%; Pred. No. 6.8;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 72 THQGHGHHHHGH 85

RESULT 12
ID 09K180 PRELIMINARY; PRT; 121 AA.
AC 09K180;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ZIM69;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238895; AAF86210.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.
DR PRINTS: PR00334; KININOGEN.
FT NON_TER 121
SQ SEQUENCE 121 AA; 14269 MW; 85F5E911130CFC72 CRC64;

Query Match 40.2%; Score 45; DB 2; Length 121;
Best Local Similarity 42.9%; Pred. No. 7;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 2 SNRAHGKHETGYH 15
DB 75 THQGHGHHHHGH 88

RESULT 13
ID 09K179 PRELIMINARY; PRT; 124 AA.
AC 09K179;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VRB (Fragment).
GN VRB.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BA1015;
RX MEDLINE=20327579; PubMed=10869077;
RA Schupp J.M., Klevytska A.M., Zinser G., Price L.B., Keim P.;
RT "vrB, a hypervariable open reading frame in bacillus anthracis.";
RL J. Bacteriol. 182:3989-3997(2000).
DR EMBL: AF238896; AAF86211.1; -.
DR HSSP: P13231; IHCE.
DR InterPro: IPR002395; KININOGEN.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:08:30 ; Search time 31.0444 Seconds
(without alignments)
393.277 Million cell updates/sec

Title: US-09-834-794a-1

Perfect score: 677

Sequence: 1 MOORGAIYALAVCALHMS.....RNSNRAHQKHEHYGKTPY 127

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	10.3	2027	2 S60123	hypothetical prote
2	70	10.3	2056	2 G88564	protein R10E11.1 l
3	68.5	10.1	88	2 S44999	hypothetical prote
4	68.5	10.1	1216	2 H85023	hypothetical prote
5	68.5	10.1	1398	2 H71606	hypothetical prote
6	68	10.0	120	2 JE0177	lymphocyte and mon
7	67	9.9	548	2 T49948	hypothetical prote
8	67	9.9	735	2 T45059	hypothetical prote
9	67	9.9	847	2 T45059	desmocollin 3b pre
10	67	9.9	901	1 T45059	desmocollin 3a pre
11	66	9.7	91	1 A28815	monocyte chemotatr
12	66	9.7	1046	2 S67786	hypothetical prote
13	66	9.7	2111	2 A70668	mycocerosate synth
14	66	9.7	2116	2 C86926	probable mycoceros
15	65.5	9.7	555	2 AG1914	carbon dioxide con
16	65.5	9.7	1576	2 T21172	hypothetical prote
17	65	9.6	588	2 T33815	hypothetical prote
18	64.5	9.5	197	2 S35252	proline-rich prote
19	64.5	9.5	289	2 S20888	phospholipase C (E
20	64.5	9.5	292	2 T50710	hypothetical prote
21	64.5	9.5	694	2 G95314	probable site-spec
22	64.5	9.5	858	2 B84780	starch branching e
23	64.5	9.5	1154	2 S69206	regulator protein
24	64.5	9.5	1361	2 T30884	neural specific DN
25	64	9.5	242	2 T32132	hypothetical prote
26	64	9.5	727	2 A55628	translational initia
27	63.5	9.4	133	2 AH3392	hypothetical cytos
28	63.5	9.4	289	2 AF1100	phospholipase C (l
29	63.5	9.4	295	1 F64085	probable site-spec

30	63.5	9.4	719	2 C88216	protein B0495.2 [l
31	63	9.3	143	2 T12245	ABA stress ripenin
32	63	9.3	285	2 A64217	ribosomal protein
33	62.5	9.2	198	2 E89008	protein W08A12.3 [
34	62.5	9.2	254	2 AH3219	conserved hypotnet
35	62.5	9.2	289	2 C43868	lecithinase - List
36	62.5	9.2	302	2 AG1360	transcription regu
37	62.5	9.2	327	2 B64799	hypothetical prote
38	62.5	9.2	380	1 JC4330	alcohol dehydrogen
39	62.5	9.2	388	2 AG3256	methionine adenosy
40	62.5	9.2	388	2 AG3256	cdm protein [lmo
41	62.5	9.2	399	2 A83354	hypothetical prote
42	62.5	9.2	431	2 T29188	hypothetical prote
43	62	9.2	126	2 A35766	platelet factor 4,
44	62	9.2	534	2 S31300	regulatory protein
45	62	9.2	667	2 T26136	hypothetical prote

ALIGNMENTS

RESULT 1
S60123
hypothetical protein R10E11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 02-Aug-2002
C:Accession: S60123; S40713
R:Ainscough, R.; Mortimore, B.
submitted to the EMBL Data Library, November 1995
A:Reference number: S60123
A:Accession: S60123
A:Molecule type: DNA
A:Residues: 1-2027 <AIN>
A:Cross-references: EMBL:Z29095; NID:9436453; PID:q1067032
A>Note: this is a revision to the sequence from reference S40713
R:Ainscough, R.; Mortimore, B.
submitted to the EMBL Data Library, December 1993
A:Reference number: S40713
A:Accession: S40713
A:Molecule type: DNA
A:Residues: 1-466, 'CKYTRRVASFISGK', 467, 'FEHFR', 474-475, 'KRLEPPKISLHSHF', 479-1986, '
A:Cross-references: EMBL:Z29095
A>Note: this sequence has been revised in reference S60123
C:Genetics:
A:Introns: 14/1: 39/3: 302/3: 424/3: 467/1: 517/1: 688/1: 1759/1: 1828/2: 1892/3: 196
C:Superfamily: transcription coactivator CREB-binding protein; bromodomain homology
F:889-946/Domain: bromodomain homology <BRO>

Query Match 10.3%; Score 70; DB 2; Length 2027;
Best Local Similarity 19.0%; Pred. No. 43;
Matches 23; Conservative 26; Mismatches 60; Indels 12; Gaps 4;

QY 13 VCAIHAASEALIPFSSCTEYSHHSIRLLERVMKRIQRADGCD-----LAAVILHV 67
Db 1561 IASLVHACQC-----RDACCRSRSCHKMKRVYQHTMCK-KRINTGCPVCKQIALACCYHA 1615

QY 68 KR--XRICVSPNHTVKQMKVQAAKXKNGKGVCHRRKHGGRNSNRAHQKHEHYGK 125
Db 1616 KHCSTDACTVPCANIKRQLEQKRSQQRADMMRRRREGLSHVGAAPPTVSNGT 1675

QY 126 P 126
Db 1676 P 1676

RESULT 2
G88564
protein R10E11.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Aug-2002
C:Accession: G88564
R:anonymous; The C. elegans Sequencing Consortium.
science 282, 2012-2018, 1998

```

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99065613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A:Accession: G88564
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2056 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA82353.1; PID:g3979836; GSPDB:GND0021; CESP:R10E11.1
C:Gene: R10E11.1
A:Map position: 3
C:Superfamily: transcription coactivator CREB-binding protein; bromodomain homology

Query Match 10.3%; Score 70; DB 2; Length 2056;
Best Local Similarity 19.0%; Pred. No. 44;
Matches 23; Conservative 26; Mismatches 60; Indels 12; Gaps 4;

QY 13 VCAALHASPALPIASSCCTEVSHHSIRLERVMCHRIORADGDCD-----LAAVILHV 67
DB 1561 IASLVHACCC---RDACRRMSCHMKRKYVOHTWCK-KRNGTCVPYCKOLIALCYHA 1615
QY 68 KR-KRICVSPNHTVQKMKVQAAXKNGKGNVCRRKHCHGCRNSNRAOGKHETVGHKT 125
DB 1616 KICSTDACTVPCNMIRQKLAEGKRSQGRADMMRRRMEIGSHVGAAPTPSVSNGT 1675
QY 126 P 126
DB 1676 P 1676

RESULT 3
S44999
Hypotheetical protein, 10K - grapevine virus A
C:Species: grapevine virus A
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 23-Mar-2001
C:Accession: S44999
R:Minifea, A.; Saldarelli, P.; Griceo, F.; Martelli, G.P.
submitted to the EMBL Data Library, October 1993
A:Description: Nucleotide sequence of the 3' terminal region of the RNA of two filamentous
A:Reference number: S44997
A:Accession: S44999
A:Molecule type: DNA
A:Residues: 1-88 <MIN>
A:Cross-references: EMBL:X75433; MID:g1405614; PIDN:CAA53186.1; PID:g468391

Query Match 10.1%; Score 68.5; DB 2; Length 88;
Best Local Similarity 27.1%; Pred. No. 3.7;
Matches 19; Conservative 14; Mismatches 20; Indels 17; Gaps 4;

QY 34 VSHNISR--RL--BRVNMCR--IQRADGDCDLAAVILH-----VKRRKICVSP 76
DB 1 MTHNHSRVGRLLSVGARBMNVCKGCAIMHNKDKCKSSISSHKLDRLRFVKEGRVALTG 60
QY 77 HNHVYKQMK 86
DB 61 ETPVYRTWVK 70

RESULT 4
H85023
Hypotheetical protein ATAg901860 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85023
R:anonymous. The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 768-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85023
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1216 <STO>

```

```

A:Cross-references: GB:NC_001268; NID:g7268570; PIDN:CAB80679.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g01860
A:Position: 4

Query Match      10.1%; Score 68.5; DB 2; Length 1216;
Best Local Similarity 23.6%; Pred. No. 39;
Matches 21; Conservative 13; Mismatches 38; Indels 17; Gaps 3;

OY 26 IASSCCTEVSHHISRLLERYNMCRIQADGDCDLAAVILHVKRRICVSPHNTVKQM 85
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 782 VASVCVSNMNMHMSDVPNLPDMCDQDAVADDCESPRLILISGAKRV-----VTSWL 833
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 86 KVAAXKNG---KGVNCHRRKKHHGRNS 110
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 834 -----LRNGRRKKKGCESCISDNGHNRASS 857
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
H71606
hypothetical protein PF80755w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: H71606
R:Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Arawind, L.; Koonin, E.;
  : Partea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.
  : Science 282, 1126-1132, 1998
A:title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: H71606
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1398 <GAR>
A:Cross-references: GB:AE001416; GB:AE001362; NID:g3845268; PIDN:MAC71940.1; PID:g384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PF80755w

Query Match      10.1%; Score 68.5; DB 2; Length 1398;
Best Local Similarity 22.6%; Pred. No. 45;
Matches 21; Conservative 17; Mismatches 48; Indels 7; Gaps 2;

OY 34 VSH-HISRLLERYNMCRIQADGDCDLAAVILHVKRRICVSPH-----NHTVKQMK 86
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1047 VKHIPQVYKSKFRRTNCKMDNRKKDLSLTIKKKEKKRIHDHINGESYVNVSKGS 1106
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 87 VQAAXKNGKGVNCHRRKKHHGRNSRAHQKHE 119
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1107 IPVMLKRLNVREKEHLKKNEKEDCSKDE 1139
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 6
JE0177
lymphocyte and monocyte chemottractant CC chemokine - human
C:Species: Homo sapiens (man)
C:Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 17-Mar-1999
C:Accession: JE0177
R:Youn, B.S.; Zhang, S.; Broxmeyer, H.E.; Antol, K.; Fraser Jr., M.J.; Hengoc, G.; K
  : Blochem. Biophys. Res. Commun. 247, 217-222, 1998
A:title: Isolation and characterisation of LMC, a novel lymphocyte and monocyte chemo
  : A:Reference number: JE0177; MUID:98308096; PMID:9642106
A:Accession: JE0177
A:Molecule type: mRNA
A:Residues: 1-120 <YOU>

Query Match      10.0%; Score 68; DB 2; Length 120;
Best Local Similarity 24.1%; Pred. No. 5.5;
Matches 21; Conservative 21; Mismatches 29; Indels 16; Gaps 4;

OY 8 IVALAVCAALIASAIIIP-----IASSCCTEVSHHISRLLERYNMCRI---QADGCDL 60
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 11 LVLLIITISARSQPKYEVWNTPTSCCL-----KYIEKVPRLRLVYGYKRALNCHL 62
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```



```

A:Experimental source: strain 5288C
A:Genetics:
A:Gene: MIPS:YDL223c
A:Cross-references: SGD:S0002382
A:Map position: 4L

Query Match
Best Local Similarity 9.7%: Score 66; DB 2; Length 1046;
Matches 17; Conservative 6; Mismatches 25; Indels 14; Gaps 2;

OY 75 SPHNHTVQMMKVQAAKH-NGKGNVCRRKKHKKRNSRRA-----HGGKHET 120
Db 943 SIHNHTTSGQKRVSVSGMSGSKSKHHHHHHHRSKRNSSKSGSDYDNNSTHSEHTPRHHQ 1002
OY 121 YG 122
Db 1003 YG 1004

RESULT 13
A70668
mycocerosate synthase (EC 2.3.1.111) - Mycobacterium tuberculosis
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70668
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
R:Comor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70668
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2111 <COL>
A:Cross-references: GB:283858; GB:AL123456; NID:g3261675; PIDN:CAB06108.1; PID:g3261616
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: mas
C:Superfamily: mycocerosic acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I
name homology: [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: acyltransferase; carrier protein; coenzyme A; phosphotransferase
E:27-426/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
E:534-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
E:1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>
E:1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>
F:2016-2096/Domain: acyl carrier protein homology <ACPr>

Query Match
Best Local Similarity 9.7%: Score 66; DB 2; Length 2111;
Matches 20; Conservative 5; Mismatches 27; Indels 6; Gaps 2;

OY 5 GLAYALAVCAALHASEALPEIASCCTEVSHHS-----RRLERVMKCIQRAGD 57
Db 180 GLAVVHLA-CRSLHDGEDALAGCAVLEPHASVNASAGMGLSTGRCHSFADAD 236

RESULT 14
C86926
probable mycocerosic synthase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 24-May-2001
C:Accession: C86926
R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R: Davies, R.M.; Devlin, K.; Dutroy, S.; Feltham, T.; Fraser, A.; Hamlin, N.; Holtr
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A:Title: Massive gene decay in the leprosy bacillus
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C86926
A:Status: preliminary

```

```

A:Experimental source: strain 5288C
A:Genetics:
A:Gene: MIPS:YDL223c
A:Cross-references: SGD:S0002382
A:Map position: 4L

Query Match
Best Local Similarity 9.7%: Score 66; DB 2; Length 1046;
Matches 17; Conservative 6; Mismatches 25; Indels 14; Gaps 2;

OY 75 SPHNHTVQMMKVQAAKH-NGKGNVCRRKKHHKGRNSRRA-----HGGKHET 120
Db 943 SIHNHTTSGQKRVSVSGMSGCKSKHHHHHHHRSKRNSSKSGDYDYNNSTHSEHTPRHHQ 1002
OY 121 YG 122
Db 1003 YG 1004

RESULT 13
A70668
mycocerosate synthase (EC 2.3.1.111) - Mycobacterium tuberculosis
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70668
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
R:Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70668
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2111 <COL>
A:Cross-references: GB:283858; GB:AL123456; NID:g3261675; PIDN:CAB06108.1; PID:g3261616
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: mas
C:Superfamily: mycocerosic acid synthase; 3-oxoacyl-[acyl-carrier-protein] synthase I
name homology: [acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: acyltransferase; carrier protein; coenzyme A; phosphotransferase
E:27-426/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <ONS>
E:534-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
E:1444-1733/Domain: long-chain alcohol dehydrogenase homology <LADH>
E:1765-1945/Domain: short-chain alcohol dehydrogenase homology <SADH>
F:2016-2096/Domain: acyl carrier protein homology <ACPr>

Query Match
Best Local Similarity 9.7%: Score 66; DB 2; Length 2111;
Matches 20; Conservative 5; Mismatches 27; Indels 6; Gaps 2;

OY 5 GLAYALAVCAALHASEALPEIASCCTEVSHHS-----RRLERVMKCIQRAGD 57
Db 180 GIMAVHIA-CRSLHDGEDALAGCAVLEPHIASVNASAGMGLSTGRCHSFADAD 236

RESULT 14
C86926
probable mycocerosic synthase [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 24-May-2001
C:Accession: C86926
R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R: Davies, R.M.; Devlin, K.; Dutroy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holtr
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A:Title: Massive gene decay in the leprosy bacillus
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C86926
A:Status: preliminary

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:07:25 ; Search time 64.2056 Seconds
(without alignments) 407.566 Million cell updates/sec

Title: US-09-834-794a-1

Perfect score: 677
Sequence: 1 MOORGIAIYALVALCAALHAS.....RNSNRAHQKHETGYHKTPY 127

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_RHEML_21:*
2: SP_Archea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Minc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_Virus:*
16: SP_Bacteriophage:*
17: SP_Archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431	63.7	135	11	O91Y39
2	162.5	24.0	120	11	O9DAU6
3	158	23.3	135	11	O9DAZ4
4	104.5	15.4	93	13	O9PTF8
5	96	14.2	144	11	O9QYV6
6	83	12.3	851	10	O94GE4
7	80.5	11.9	89	13	O91BE0
8	77.5	11.4	91	13	O8G556
9	76	11.2	101	13	O93238
10	71.5	10.6	90	13	O9PW46
11	71.5	10.6	90	13	O91OC9
12	71	10.5	584	10	O8RUT3
13	69.5	10.3	633	4	O96L88
14	69.5	10.3	1020	4	O9C0D4
15	69.5	10.3	1468	5	O9V854
16	69	10.2	151	11	O9D604

17	68.5	10.1	88	12	O67708	O67708 grapevine v
18	68.5	10.1	363	4	O96P69	O96P69 homo sapien
19	68.5	10.1	815	3	O9HER3	O9HER3 neurospora
20	68.5	10.1	1216	10	O9SY14	O9SY14 arabidopsis
21	68.5	10.1	1308	10	O8RXC4	O8RXC4 arabidopsis
22	68.5	10.1	1398	5	O96244	O96244 plasmidium
23	68	10.0	100	3	O8Q555	O8Q555 gallus gall
24	67	9.9	118	12	O9J2M1	O9J2M1 macaca mula
25	67	9.9	197	16	O9RJ59	O9RJ59 streptomyce
26	67	9.9	446	5	O9VJ92	O9VJ92 drosophila
27	67	9.9	548	10	O9LXG8	O9LXG8 arabidopsis
28	67	9.9	735	5	O9NE57	O9NE57 caenorhabd
29	66	9.7	95	13	O8QCV9	O8QCV9 paratichthy
30	66	9.7	222	10	O8S1H5	O8S1H5 oryza sativ
31	66	9.7	224	5	O95PY7	O95PY7 caenorhabd
32	66	9.7	1046	3	O07653	O07653 saccharomyc
33	66	9.7	1300	4	O15090	O15090 homo sapien
34	66	9.7	2111	16	P96291	P96291 mycobacteri
35	66	9.7	2116	16	O9CD78	O9CD78 mycobacteri
36	65.5	9.7	294	2	O9K187	O9K187 bacillus ce
37	65.5	9.7	555	16	O8XY13	O8XY13 anabaena sp
38	65.5	9.7	883	5	O8SOJ7	O8SOJ7 encaphalito
39	65.5	9.7	1576	5	P90841	P90841 caenorhabd
40	65	9.6	110	10	O82575	O82575 lycopersico
41	65	9.6	216	10	O9ST96	O9ST96 oryza sativ
42	65	9.6	508	5	O9TV04	O9TV04 caenorhabd
43	65	9.6	607	5	O9VHA9	O9VHA9 drosophila
44	64.5	9.5	97	11	O92318	O92318 cavia porce
45	64.5	9.5	104	12	O41924	O41924 molluscum c

ALIGNMENTS

RESULT 1
ID O91Y39 PRELIMINARY; PRT; 135 AA.
AC O91Y39;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DR 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
PT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE CC chemokine CCL28.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
NC Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary Gland;
RA Beuken E., Gruljthuisen Y.K., Bruggeman C.A., Vink C.;
RT "Rattus norvegicus CC chemokine CCL28 (rcccl28) mRNA.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF361490; AAK52773.1;
DR InterPro; IPR001811; Chemokine_IL8.
SQ SEQUENCE 135 AA; 15096 MW; AFF79E0EF7EA8A64 CRC64;

Query Match 63.7%; Score 431; DB 11; Length 135;
Best Local Similarity 63.3%; Pred. No. 1.1e-43;
Matches 88; Conservative 9; Mismatches 24; Indels 18; Gaps 4;
OY 1 MOORGIAIYALVALCAALHASAIPDIASSCCTEVSHHSIRLLERVNMCRIGADGDDL 60
DB 1 MOOAGLFLVAVVAVCAFTSEALIPDIASSCCTEVSHHPRLLERVNSCSIGRADGDDL 60
OY 61 AAVIHLVRRKRCVSPHNITVQMKVQAAXKNGNGNCRKKH-----HGK 107
DB 61 AAVIHLVRRKRCVSPHNITVQMKVQAAXKNGNGNCRKKH-----HGK 107
OY 108 RNSNRAHQKHETGYHKTP 126
DB 120 HCTRRIH-GTHD---HEAP 134

```

RESULT 2
Q9DAU6 PRELIMINARY: PRT: 120 AA.
AC Q9DAU6:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Small Inducible cytokine A27.
GN SCYA27.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glassl C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schriml L. M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Furuno M., Aono H., Balderelli R., Barsch G.,
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamuya M., Lee N. H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. -F.,
RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS GLAND;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK005520; BAB24095.1; -
DR EMBL: BC028511; AAH28511.1; -
DR MGD: MGI:1343459; Scya27.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00199; SCY; 1.
SQ SEQUENCE 120 AA; 13464 MW; D1EEEL270AB580BF CRC64;

Query Match 24.0%; Score 162.5; DB 11; Length 120;
Best Local Similarity 41.8%; Pred. No. 1.2e-11;
Matches 33; Conservative 14; Mismatches 29; Indels 3; Gaps 2;

OY 19 ASFAITPIAS--SCREVSHH-ISRRLERVMNCRIGRADGDCDLAVILHVKRXCIVS 75
DB 21 AFAALPLPSSSCCTQLRQPLPSRLRLRIVMELEADGDCDLQAVVILHARRSVCVA 80
OY 76 PHNHTVKOMMKVQAAXKNG 94
DB 81 PQRSLARWLERQGRKRLQG 99

RESULT 3
Q9DAZ4 PRELIMINARY: PRT: 135 AA.
AC Q9DAZ4:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Small Inducible cytokine A27.
GN SCYA27.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glassl C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schriml L. M., Staudil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Furuno M., Aono H., Balderelli R., Barsch G.,
RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamuya M., Lee N. H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. -F.,
RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
RX EMBL: AK005398; BAB24001.1; -
DR MGD: MGI:1343459; Scya27.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00199; SCY; 1.
SQ SEQUENCE 135 AA; 15751 MW; 956BBE984B2E389 CRC64;

Query Match 23.3%; Score 158; DB 11; Length 135;
Best Local Similarity 40.8%; Pred. No. 4.7e-11;
Matches 31; Conservative 16; Mismatches 27; Indels 2; Gaps 2;

OY 21 EAI-LPIASCTEVSHH-ISRRLERVMNCRIGRADGDCDLAVILHVKRXCIVS 78
DB 39 EALPLPSSSCCTQLRQPLPSRLRLRIVMELEADGDCDLQAVVILHARRSVCVHQ 98
OY 79 HTVKOMMKVQAAXKNG 94
DB 99 RSLARWLERQGRKRLQG 114

RESULT 4
Q9PTF8 PRELIMINARY: PRT: 93 AA.
AC Q9PTF8:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CC chemokine CCL1.
GN CCL1 OR CCL1.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Tseng S., Rollins B. J.;
RT *Chemokines and chemokine receptors in zebrafish.*;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF201450; AAF17560.1; -
DR ZFIN: ZDB-GENE-000208-28; ccl1.
DR InterPro: IPR001811; Chemokine_IL8.
DR Pfam: PF00048; IL8; 1.
DR SMART: SM00199; SCY; 1.
SQ SEQUENCE 93 AA; 10776 MW; 17C9B5309903ED7 CRC64;

Query Match 15.4%; Score 104.5; DB 13; Length 93;
Best Local Similarity 31.0%; Pred. No. 7.7e-05;

```


Matches 22: Conservative 11: Mismatches 35: Indels 3: Gaps 1:

OY 6 LAIYALVCAALHASEAILPIASSCCTEVSHHSRLLERVMNCRIQADGCDLAAVL 65
DB 9 LLLVCFITIIITLDMKGAIF---TCCLSVLRIRIPKRVLRVRRVEVDQTSCHCEIKALIL 65

OY 66 HVKRRICVSP 76
DB 66 HFKGKRICAMP 76

RESULT 5

OY 090Y6 PRELIMINARY; PRT; 144 AA.
AC 090Y6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Thymus-expressed chemokine precursor.
GN SCYA25 OR TECK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20069400; PubMed=10602049;
RA Wudel M.A., Philippe J.M., Nguyen C., Victorero G., Freeman T.,
RA Wooding P., Mizek A., Mattel M.G., Malsen M., Jordan B.R.,
RA Malsen B., Carrier A., Naquet P.;
RT "The chemokine TECK is expressed by thymic and intestinal epithelial
RT cells and attracts double- and single-positive thymocytes expressing
RT the TECK receptor CCR9.";
RL Eur. J. Immunol. 30:262-271(2000).
DR EMBL: AJ249480; CAB5762.1; -.
DR MGD: MGI:1099448; Scya25.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: SM00199; SCY; 1.
KW Signal.
FT CHAIN 1 23 POTENTIAL.
FT SIGNAL 24 144 THYMUS-EXPRESSED CHEMOKINE.
SO SEQUENCE 144 AA; 16732 MW; CEC1924B320B4408 CRC64;

Query Match 14.2%; Score 96; DB 11; Length 144;
Best Local Similarity 25.0%; Pred. No. 0.0013;
Matches 32; Conservative 19; Mismatches 61; Indels 16; Gaps 4;

OY 8 IVALAVCA---ALHASEAILPIASSCCTEVSHHSRLLERVMNCRIQADGCDLAAVL 64
DB 9 LVACVGMMPVYHAGAF---EDCCGQYHRIKMNVLRRARNYHQDEVSGSCNLRVAVR 64

OY 65 LHVXRKRICVSPHNHTVYKQMKVQAA-----KNGKGNVCHRRKHH--GKRNENRAHQ 115
DB 65 FYFQKQVCGNPEDMNVRARILTKRLVYHKSASDSQTERKKSNNMKSKVENPNST 124

OY 116 GKHEYYGH 123
DB 125 VRSATLGH 132

RESULT 6

OY 094GE4 PRELIMINARY; PRT; 851 AA.
AC 094GE4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative dext protein.
GN OJ111_B11.16.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euharidiales; Oryzaceae; Oryza.

OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Bueli C.R., Yuan Q., Ouyang S., Moffet K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shwartsbeyn M., Tstlin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pal G., Vanaken S.E.,
RA Salzbach T.R., Feldlyum T.V., Kalb E., Quackenbush J.,
RA Salzbach S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OJ111_B11 genomic sequence.";
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC091247; AAK02448.1; -.
SO SEQUENCE 851 AA; 94842 MW; C9F90276AF0E208 CRC64;

Query Match 12.3%; Score 83; DB 10; Length 851;
Best Local Similarity 22.9%; Pred. No. 0.33;
Matches 35; Conservative 20; Mismatches 50; Indels 48; Gaps 9;

OY 2 QORGLAIYALVCAALHASEAILPIASSCCTEV-----SHHSRLLERVMNCRIQAD 55
DB 613 KSKGLTATTSFPGYLYIE-----GSSGACADVVDIGETSY-----SMVLADNV 657

OY 56 GDCDLAAVILHVKKRIRICV---SPHNHTVYKQMKVQAAKNGKGNVCHRRKHH-----HG 106
DB 658 GDDDLIVITWNGNVCFSTSPSPH-HPLKEW---RSSNQRNNAAYRYNREGIYVKG 712

OY 107 KRNENRAHQKH-----ETVGHKTPY 127
DB 713 SR-TRDEGKHFVFEIVDKYRVPYGNQAPY 744

RESULT 7

OY 091BE0 PRELIMINARY; PRT; 89 AA.
AC 091BE0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Chemokine K203 precursor.
GN K203.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20170941; PubMed=10704244;
RA Sick C., Schneider K., Staeheli P., Welting K.C.;
RT "Novel chicken CXK and CC chemokines.";
RL Cytokine 12:181-186(2000).
DR EMBL: Y18692; CAB70956.1; -.
DR HSSP: P13236; IHUM.
DR InterPro: IPR001811; Chemokine_IL8.
DR SMART: PF00048; IL8; 1.
DR SMART: SM00199; SCY; 1.
KW Signal.
FT CHAIN 1 21 POTENTIAL.
FT SIGNAL 22 89 CHEMOKINE K203.
SO SEQUENCE 89 AA; 9896 MW; 6FAZEA7A4950CA75 CRC64;

Query Match 11.9%; Score 80.5; DB 13; Length 89;
Best Local Similarity 27.6%; Pred. No. 0.054;
Matches 24; Conservative 18; Mismatches 34; Indels 11; Gaps 4;

OY 6 LAIYALVCAALHASEA---ILPIASSCCTEVSHHSRLLERVMNCRIQADGCDL 60
DB 3 LSAVVALILIASFCRSASAPGDPVPTCTTYIRHKIPRMLIQ-----RHVSTSCSK 57

OY 61 AAVILHVKKRIR-ICVSPHNHTVYKQMK 86
DB 58 PALITTKREVCANPSPDPVQRYLQ 84

RESULT 8
 080G56 PRELIMINARY; PRT; 91 AA.
 AC 080G56;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Chemokine ah221.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21655115; PubMed=11797102;
 RA Hughes S., Haynes A., O'Regan M., Bumstead N.;
 RT "Identification, mapping, and phylogenetic analysis of three novel
 chicken CC chemokines.";
 RL Immunogenetics 53:674-683(2001).
 DR EMBL: AY037860; AAK84433.1; -
 SO SEQUENCE 91 AA; 9898 MW; 27D489DC9E51D075 CRC64;

Query Match 11.4%; Score 77.5; DB 13; Length 91;
 Best Local Similarity 29.2%; Pred. No. 0.13;
 Matches 26; Conservative 19; Mismatches 35; Indels 9; Gaps 4;
 QY 7 ATVALAVCAALHASEALIPASCC--TEVSHHSRLERVMNCRIQRADGCDLAAVIL 65
 DB 10 ALALLMDCSSAVNHDLGP--TTCCFSYORPPRNLIASAVI-----TSSKRLPAVIL 62
 QY 66 HVKRXR-ICVSPHNHTVKOMKVOAAXKN 93
 DB 63 VFKKREICVNPESVWOKRLELLQKQEN 91

RESULT 9
 093238 PRELIMINARY; PRT; 101 AA.
 AC 093238;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CC chemokine-1.
 OS Cyprinus carpio (Common carp).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Cyprinus.
 OX NCBI_TaxID=7962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99367403; PubMed=10436187;
 RA Fujiki K., Shin D., Nakao M., Yano T.;
 RT "Molecular cloning of carp (Cyprinus carpio) CC chemokine, CXK
 chemokine receptors, allograft inflammatory factor-1, and natural
 killer cell enhancing factor by use of suppression subtractive
 hybridization.";
 RL Immunogenetics 49:909-914(1999).
 DR EMBL: AB010469; BAA31459.1; -
 DR HSSP: P13500; IMCA.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR SMART: SM00199; SCY; 1.
 SO SEQUENCE 101 AA; 11266 MW; 384B648FD5CBA68 CRC64;

Query Match 11.2%; Score 76; DB 13; Length 101;
 Best Local Similarity 26.5%; Pred. No. 0.22;
 Matches 22; Conservative 15; Mismatches 42; Indels 4; Gaps 2;
 QY 4 RGLATVALAVCAALHASEALIPASCCTEVSHHSRLERVMNCRIQRADGCDLAAV 63
 DB 9 RSLAAVAVIASVITWTADTAIVYSCCTKVT--TAETVDPLINTIRLORESLPC-VKAV 64

QY 64 ILHVKKRXRICVSPHNHTVKOMK 86
 DB 65 IFKTKGECSDPKLRWVKEKVK 87

RESULT 10
 09PWA6 PRELIMINARY; PRT; 90 AA.
 AC 09PWA6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Chemokine.
 GN SCY4.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hughes S.M., Bumstead N.;
 RT "Mapping of the gene encoding the chicken homologue of the mammalian
 chemokine SCY4.";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF146730; AAD48772.1; -
 DR HSSP: P13236; 1HUM.
 DR InterPro: IPR000827; CC_chemokine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; 1.
 SO SEQUENCE 90 AA; 9986 MW; 50AF9679A26751CB CRC64;

Query Match 10.6%; Score 71.5; DB 13; Length 90;
 Best Local Similarity 27.1%; Pred. No. 0.66;
 Matches 23; Conservative 17; Mismatches 32; Indels 13; Gaps 5;
 QY 6 LATVALAVCAALHASEALIPAS---SCTEVSHHSRLERVMNCRIQRADGCDLA 61
 DB 8 LAVLLAICTQYSNA----PVGSDPTSCFT--YISHQLPSP-VADYETTSQCPHA 59
 QY 62 AVILHVKKRXR-ICVSPHNHTVKOM 85
 DB 60 GVEITRKREVCANPNQNDVQDYM 84

RESULT 11
 0910C9 PRELIMINARY; PRT; 90 AA.
 AC 0910C9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Macrophage inflammatory protein 1-beta.
 GN SCY4.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hughes S.M., Bumstead N.;
 RT "Mapping of the gene encoding the chicken homologue of the mammalian
 chemokine SCY4.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ243034; CAB45103.1; -
 DR InterPro: IPR000827; CC_chemokine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR Pfam: PF00048; IL8; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; UNKNOWN_1.

SO SEQUENCE 90 AA; 9987 MW; 50AF9679A267408F CRC64;

Query Match 10.6%; Score 71.5; DB 13; Length 90;
Best Local Similarity 27.1%; Pred. No. 0.66;
Matches 23; Conservative 17; Mismatches 32; Indels 13; Gaps 5;

OY 6 LAVALAVALCAALHASEALIPLAS---SCTEVSHHSRLRLERVNMCRIORADGCDLA 61
DB 8 LAVLIALACVOTSA---PVGSDPPTSCTP---YISRQLPFSF-VADYETNSQCPHA 59

OY 62 AVILHVKRXR-ICVSPHNHTVKOMW 85
DB 60 GVETTRKREVCANPENDWODYM 84

RESULT 12

OBRUT3 PRELIMINARY; PRT; 584 AA.

AC OBRUT3; PRELIMINARY; PRT; 584 AA.

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

GN P0703B11.13 OR P0485B12.5

OS Oryza sativa (Japanese cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Eubacteroidae; Oryzae; Oryza.

NCBI_TaxID=39947;

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP003302; BAB85294.1;

DR EMBL; AP003348; BAB86475.1;

SO SEQUENCE 584 AA; 65103 MW; 835045ACD9B8CD29 CRC64;

Query Match 10.5%; Score 71; DB 10; Length 584;

Best Local Similarity 26.5%; Pred. No. 5.9;

Matches 30; Conservative 13; Mismatches 50; Indels 20; Gaps 5;

OY 3 ORGLAVALAVALCAALHASEALIPLAS---SCTEVSHHSRLRLERVNMCRIORADGCD 57

DB 274 ERGLEFRACACGIVGHASCLP---TCRCGETTHYSDACPLKRVTFLEGDHVPKD 330

OY 58 CDLAVALHVKRXR-----ICVSPHNHTVKOMKVOAAXKXGKGV 98

DB 331 COLNAVIAIKTKKEQGTVOPIROPMTIDNSGHNPSALPPTAPVYANHGKRSNV 383

RESULT 13

O96LNB PRELIMINARY; PRT; 633 AA.

AC O96LNB; PRELIMINARY; PRT; 633 AA.

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK058072; BAB71650.1;

DR EMBL; AK058072; BAB71650.1;

DR InterPro; IPR000822; Znf_C2H2.

DR Pfam; PF00096; Zf-C2H2.1.

DR SMART; SM00355; Znf_C2H2.1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 1.

DR DNA-binding; zinc-finger.

FT NON_TER

SO SEQUENCE 1020 AA; 113717 MW; C28BCDF143856177 CRC64;

RP SEQUENCE FROM N.A.

RC TISSUE=TESTIS;

RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,

RA Horita T., Hiraka S., Murekawa K., Takiguchi S., Kusano J., Chiba Y.,

RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,

RA Suzuki Y., Hata H., Nagawa K., Mizuno S., Morinaga M., Kawamura M.,

RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,

RA Kawakami B., Nagai K., Isogai T., Sugano S.;

RA "NEDO human cDNA sequencing project."

RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AK058072; BAB71650.1;

DR InterPro; IPR000822; Znf_C2H2.

DR Pfam; PF00096; Zf-C2H2.1.

DR SMART; SM00355; Znf_C2H2.1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 1.

DR DNA-binding; zinc-finger.

FT NON_TER

SO SEQUENCE 633 AA; 70189 MW; DE5ED216E0606E03 CRC64;

Query Match 10.3%; Score 69.5; DB 4; Length 633;

Best Local Similarity 28.9%; Pred. No. 9.8;

Matches 33; Conservative 25; Mismatches 37; Indels 19; Gaps 7;

OY 12 AVCAALHASEALIPLASCTEVSHHS-RRLERVNMCRIORADGCDLAVALHVKRX 70

DB 335 AVLRVLSSENAHIEATCEAPVSPICSERLIKRPVPCPVQADSDIQ---PLRSRG 390

OY 71 RIVSPHNHT-VKQMKVOAAXKXGKGV--HKKH--HGKRSNAH-OGK 117

DB 391 PIDMSPNIETPLRLKRESA-----VCSTIHKRTGLYGOOGSSSELNKG 437

RESULT 14

O9C0D4 PRELIMINARY; PRT; 1020 AA.

AC O9C0D4; PRELIMINARY; PRT; 1020 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

GN KIAA1729 protein (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;

RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB051516; BAB21820.1;

DR EMBL; AB051516; BAB21820.1;

DR InterPro; IPR000822; Znf_C2H2.

DR Pfam; PF00096; Zf-C2H2.1.

DR SMART; SM00355; Znf_C2H2.1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.

DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 1.

DR DNA-binding; zinc-finger.

FT NON_TER

SO SEQUENCE 1020 AA; 113717 MW; C28BCDF143856177 CRC64;

Query Match 10.3%; Score 69.5; DB 4; Length 1020;

Best Local Similarity 28.9%; Pred. No. 17;

Matches 33; Conservative 25; Mismatches 37; Indels 19; Gaps 7;

OY 12 AVCAALHASEALIPLASCTEVSHHS-RRLERVNMCRIORADGCDLAVALHVKRX 70

DB 722 AVLRVLSSENAHIEATCEAPVSPICSERLIKRPVPCPVQADSDIQ---PLRSRG 777

OY 71 RIVSPHNHT-VKQMKVOAAXKXGKGV--HKKH--HGKRSNAH-OGK 117

DB 778 PIDMSPNIETPLRLKRESA-----VCSTIHKRTGLYGOOGSSSELNKG 824

Search completed: January 14, 2003, 18:15:26
Job time : 67.2056 secs

```
RESULT 15
O9V854 PRELIMINARY; PRT; 1468 AA.
ID O9V854
AC O9V854;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CG10936 protein.
GN CG10936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman I.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan N., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003802; AAF57823.1; -
DR FlyBase; FBgn0034253; CG10936.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF01390; SEA; 1.
SQ
SEQUENCE 1468 AA; 160314 MW; 8BA6D6B65BFA5184 CRC64;
```

Query Match 10.3%; Score 69.5; DB 5; Length 1468;
Best Local Similarity 28.8%; Pred No. 25;
Matches 17; Conservative 9; Mismatches 24; Indels 9; Gaps 2;

OY 77 HHHTYKQMKKVOAAKXGKGNVCHRRK-HHGKRNNSRAHQKH-----ETYGCHTP 126
DB 1381 HHHTSGDYFYDDQPTYSQGDGLKRPQLHNGNNSHREYPAHHHNOASQHQAYGCHQYP 1439

RESULT 12

T34206

hypothetical protein F10E7.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34206

R:Pauley, A.

A:Submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F10E7.

A:Reference number: 221489

A:Accession: T34206

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-929 <PAU>

A:Cross-references: EMBL:U41264; PIDN:AAA82421.1; CESP:F10E7.8

C:Genetics:

A:Gene: CESP:F10E7.8

A:Introns: 13/1; 70/3; 180/3; 672/3; 701/1; 855/1

Query Match

Best Local Similarity 39.3%; Score 44; DB 2; Length 929;

Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 NSNRHQGHETTYGKTP 18

DB 673 NTNKLQSRHEIYAVNYP 690

RESULT 13

T31084

RNA-directed DNA polymerase (EC 2.7.7.49) - common earwig non-LTR retrotransposon R2 (fr

N:Alternate names: reverse transcriptase

C:Species: Forficula auricularia (common earwig)

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T31084

R:Burke, W.D.; Malik, H.S.; Jones, J.P.; Elckbush, T.H.

Mol. Biol. Evol. 16, 502-511, 1999

A:Title: The domain structure and retrotransposition mechanism of R2 elements are conserved

A:Reference number: 220937; MUID:99261661; PMID:10331276

A:Accession: T31084

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1033 <BUR>

A:Cross-references: EMBL:AF015819; NID:g3559775; PID:g3559776; PIDN:AAC34906.1

C:Genetics:

A:Mobile element: non-LTR retrotransposon R2

C:Keywords: nucleotidyltransferase

Query Match

Best Local Similarity 39.3%; Score 44; DB 2; Length 1033;

Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 NSNRHQGHETTYG 15

DB 937 NDEAHQAKAKERYAH 951

RESULT 14

S14431

knob-associated histidine-rich protein - malaria parasite (Plasmodium falciparum) (fragm

C:Species: Plasmodium falciparum

C:Date: 21-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 09-Jun-2000

C:Accession: S14431; S07194

R:Ardeshtir, F.

submitted to the EMBL Data Library, April 1987

A:Reference number: S14431

A:Accession: S14431

A:Molecule type: mRNA

A:Residues: 1-277 <ARD>

A:Cross-references: EMBL:Y00063; NID:g9910; PIDN:CAA68270.1; PID:g939959

R:Ardeshtir, F.; Flint, J.E.; Matsumoto, Y.; Akawa, M.; Reese, R.T.; Stanley, H.

EMBO J. 6, 1421-1427, 1987

A:Title: cDNA sequence encoding a Plasmodium falciparum protein associated with knobs

A:Reference number: S07194; MUID:87275837; PMID:3301326

A:Accession: S07194

A:Molecule type: mRNA

A:Residues: 1-275 <ARD1>

A:Cross-references: EMBL:Y00063

C:Superfamily: knob-associated histidine-rich protein

C:Keywords: glycoprotein

Query Match 38.8%; Score 43.5; DB 2; Length 277;

Best Local Similarity 34.6%; Pred. No. 31;

Matches 9; Conservative 2; Mismatches 4; Indels 11; Gaps 1;

QY 5 AHQGH-----ETTYGKTPY 19

DB 93 AHQSHGNLRGHNGKSGEGYEAPY 118

RESULT 15

A54494

knob-associated histidine-rich protein precursor - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000

C:Accession: A54494

R:Ellis, J.; Irving, D.O.; Wellens, T.E.; Howard, R.J.; Cross, G.A.M.

Mol. Biochem. Parasitol. 26, 203-214, 1987

A:Title: Structure and expression of the knob-associated histidine-rich protein of P.

A:Reference number: A54494; MUID:88122262; PMID:2448620

A:Accession: A54494

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-473 <ELL>

A:Cross-references: GB:MJ9028; NID:g160359; PIDN:AAA29630.1; PID:g160360

C:Genetics:

A:Map position: 2

C:Superfamily: knob-associated histidine-rich protein

Query Match

Best Local Similarity 38.8%; Score 43.5; DB 2; Length 473;

Matches 9; Conservative 2; Mismatches 4; Indels 11; Gaps 1;

QY 5 AHQGH-----ETTYGKTPY 19

DB 278 AHQSHGNLRGHNGKSGEGYEAPY 303

Search completed: January 14, 2003, 18:16:24
Job time: 6.64444 secs

C:Keywords: glycoprotein; iron; metalloprotein; phosphoric monoester hydrolase; zinc
F:141-209/Domain: phosphoesterase core homology <PPEC>
F:85,392/Binding site: carbonylase (Asn) (covalent) #status predicted
F:147,174,177,332/Binding site: Iron (Asp, Asp, Tyr, His) #status predicted
F:114,207,291,330/Binding site: zinc or manganese (Asp, Asn, His, His) #status predicted
F:208,301/Active site: His #status predicted

Query Match 40.28; Score 45; DB 1; Length 427;
Best Local Similarity 66.78; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
OY 1 NSNRHOGKHET 12
DB 296 NSNRHOGKHET 307

RESULT 8
T02808
conserved hypothetical protein YPL199C, L2602.6 [Imported] - Leishmania major (strain F)
C:Species: Leishmania major
C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C:Accession: D81457; T02808
R:Myler, P.J.; Audleman, L.; DeVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A:Reference number: A81455; MUID:99178987; PMID:10077609
A:Accession: D81457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-897 <PVL>
A:Cross-references: GB:AE001274; NID:93264850; PIDN:AAC24631.1; PID:9295584; GSPDB:GNOC
A:Experimental source: strain MHOM/IL/81/Friedlin
C:Genetics:
A:Gene: L2602.6
A:Map position: 1

Query Match 40.28; Score 45; DB 2; Length 897;
Best Local Similarity 53.88; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 7 QGKHETYGKTPY 19
DB 847 QGKHETYGKTPY 859

RESULT 9
S02053
hypothetical protein (D10 5' region) - phage T5
C:Species: phage T5
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Oct-1999
C:Accession: S02053
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 10353-10354, 1988
A:Title: The nucleotide sequence of the region of bacteriophage T5 early genes D10-D15.
A:Reference number: S01931; MUID:89057468; PMID:30577441
A:Accession: S02053
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-165 <KAL>
A:Cross-references: EMBL:X12930; NID:915407; PIDN:CAA31397.1; PID:9579170
C:Genetics:
A:Start codon: GTG
A:Superfamily: phage T5 hypothetical protein (D10 5' region)

Query Match 39.38; Score 44; DB 2; Length 165;
Best Local Similarity 42.18; Pred. No. 15;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
OY 1 NSNRHOGKHETYGKTPY 19
DB 108 HSARITNGKCYVAKVPY 126

RESULT 10

A47398
serotonin transporter - human
N:Alternate names: 5-hydroxytryptamine transporter
C:Species: Homo sapiens (hmn)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47398; S37688
R:Ramanorthy, S.; Bauman, A.L.; Moore, K.R.; Han, H.; Yang-Feng, T.; Chang, A.S.; Ga
Proc. Natl. Acad. Sci. U.S.A. 90, 2542-2546, 1993
A:Title: Antidepressant- and cocaine-sensitive human serotonin transporter: molecular
A:Reference number: A47398; MUID:93211998; PMID:7681602
A:Accession: A47398
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-630 <RAM>
A:Cross-references: GB:L05568; NID:9291819; PIDN:AAA35492.1; PID:9291820
A:Experimental source: placental trophoblastic cell line JAR
A:Note: sequence extracted from NCBI backbone (NCBI:128282, NCBI:128283)
R:Lesch, K.P.; Wolozin, B.L.; Estler, H.C.; Murphy, D.L.; Riederer, P.
J. Neural Transm. 91, 67-72, 1993
A:Title: Isolation of a cDNA encoding the human brain serotonin transporter.
A:Reference number: S37688
A:Accession: S37688
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-630 <LES>
A:Cross-references: EMBL:X70697; NID:936432; PIDN:CAA50029.1; PID:936433
A:Note: the authors translated the codon GGA for residue 25 as Val and GAA for residu
C:Genetics:
A:Gene: GDB:SLC6A4; HTT
A:Cross-references: GDB:134713; OMIM:182138
A:Map position: 17q11.1-17q12
A:Superfamily: gamma-aminobutyric acid transporter
C:Keywords: transmembrane protein

Query Match 39.38; Score 44; DB 2; Length 630;
Best Local Similarity 50.08; Pred. No. 59;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
OY 6 HOGKHETYGKTPY 19
DB 75 HOGKHETYGKTPY 88

RESULT 11
C88397
protein H04J21.3 [Imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C88397
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99059613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: C88397
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-722 <STO>
A:Cross-references: GB:chr.III; PIDN:AAB94969.1; PID:92746805; GSPDB:GN00021; CESP:HO
C:Genetics:
A:Gene: H04J21.3
A:Map position: 3

Query Match 39.38; Score 44; DB 2; Length 722;
Best Local Similarity 61.58; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 3 NRAHOGKHETYGH 15
DB 83 NRAHOGKHETYGH 95

Db 44 HQAKRHYHSMYGNOPY 61

RESULT 3

C88216

protein B0495.2 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-May-2002

C:Accession: C88216

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998.

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:95069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: C88216

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-719 <STO>

A:Cross-references: GB:chr_II; PID:g687819; GSPDB:GN00020; CESP:B0495.2

C:Genetics:

A:Gene: B0495.2

A:Map position: 2

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match

Best Local Similarity

Matches

8; Conservative

1; Mismatches

3; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

C:Accession: S01449

R:Man'kin, A.S.; Kagramanova, V.K.

Nucleic Acids Res. 16, 4679-4692, 1988

A:Title: Complex promoter pattern of the single ribosomal RNA operon of an archaeobact

A:Reference number: S01449; MUID:88247769; PMID:3380693

A:Accession: S01449

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-170 <MAN>

A:Cross-references: EMBL:X06805

A:Note: the source is designated as Halobacterium halobium

Query Match

Best Local Similarity

Matches

8; Conservative

3; Mismatches

7; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

RESULT 6

T29339

hypothetical protein F21C10.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T29339

R:Du, Z.; Gattung, S.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid F21C10.

A:Reference number: 220610

A:Accession: T29339

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-382 <DUZ>

A:Cross-references: EMBL:U55364; PIDN:AAA97972.1; GSPDB:GN00023; CESP:F21C10.8

A:Experimental source: strain Bristol N2; clone F21C10

C:Genetics:

A:Gene: CESP:F21C10.8

A:Map position: 5

A:Introns: 20/1; 219/3; 249/3

Query Match

Best Local Similarity

Matches

8; Conservative

3; Mismatches

8; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

0; Indels

0; Gaps

RESULT 7

T49031

acid phosphatase (EC 3.1.3.2) purple F3C22.180, precursor [similarity] - Arabidopsis

N:Alternate names: purple acid phosphatase (PAP)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 14-Sep-2001

C:Accession: T49031

R:Purnelle, B.; Masny, D.; Goffeau, A.; Boutry, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.

submitted to the Protein Sequence Database, April 2000

A:Reference number: 225013

A:Accession: T49031

A:Molecule type: DNA

A:Residues: 1-427 <PUR>

A:Cross-references: EMBL:AL353912; GSPDB:GN00061; ATSP:F3C22.180

A:Experimental source: cultivar Columbia; BAC clone F3C22

A:Gene: ATSP:F3C22.180

A:Map position: 3

A:Introns: 46/3; 146/1; 227/2; 337/3; 367/2

C:Function: catalyzes the hydrolysis of phosphate monoesters

C:Superfamily: kidney bean purple acid phosphatase; phosphoesterase core homology

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:08:30 ; Search time 4.64444 Seconds
(without alignments)
393.277 Million cell updates/sec

Title: US-09-834-794A-5
Perfect score: 112
Sequence: 1 NSNRAHQKHETGYHKTPT 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR73:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	42.9	567	2	A71619
2	47	42.0	345	2	A48462
3	47	42.0	719	2	C88216
4	47	42.0	735	2	T45059
5	45	40.2	170	2	S01449
6	45	40.2	382	2	T29339
7	45	40.2	427	1	T49031
8	45	40.2	897	2	T02808
9	44	39.3	165	2	S02053
10	44	39.3	630	2	A47398
11	44	39.3	722	2	C88397
12	44	39.3	929	2	T34206
13	44	39.3	1033	2	T31084
14	43.5	38.8	277	2	S14431
15	43.5	38.8	473	2	A54494
16	43.5	38.8	634	2	A54495
17	43.5	38.8	634	2	A28412
18	43.5	38.8	654	2	B71623
19	43.5	38.8	657	2	A29454
20	43.5	38.8	155	2	E69482
21	43	38.4	555	2	T01142
22	43	38.4	792	2	A71822
23	43	38.4	964	2	D87326
24	43	38.4	991	2	S43891
25	43	38.4	1025	2	A82516
26	43	38.4	1146	2	C83304
27	42.5	37.9	513	2	T37180
28	42	37.5	102	2	AH3257
29	42	37.5	104	2	G95127

30	42	37.5	104	2	D97898	50S ribosomal prot
31	42	37.5	250	2	B35026	filaggrin B - mouse
32	42	37.5	255	2	A35026	filaggrin A - mouse
33	42	37.5	331	2	S31395	hypothetical prote
34	42	37.5	630	2	A53463	hypothetical membr
35	42	37.5	715	2	T25631	hypothetical prote
36	42	37.5	856	2	A64659	hypothetical prote
37	42	37.5	960	2	A61940	isolectin-CRNA syn
38	42	37.5	1344	2	T34188	myb-binding protei
39	41.5	37.1	335	2	S64352	dystroglycan alpha
40	41.5	37.1	650	2	S59630	dystroglycan
41	41.5	37.1	895	2	I54343	dystrophin associa
42	41.5	37.1	895	2	S20582	antifungal protei
43	41	36.6	96	2	UC4581	hypothetical prote
44	41	36.6	113	2	S57101	conserved hypothe
45	41	36.6	158	2	A83173	

ALIGNMENTS

RESULT 1
A71619
membrane transporter PFB0275W - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C/Accession: A71619
R/Gardner, M.J.; Tetteh, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Perera, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A/Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A/Reference number: A71600; MIM:6021743; PMID:9604351
A/Accession: A71619
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-567 <GAR>
A/Cross-references: GB:AE001384; GB:AE001362; NID:g3845139; PIDN:AACT1844.1; PID:g384
A/Experimental source: clone 3D7
C/Genetics:
A/Gene: PFB0275W

Query Match 42.9%; Score 48; DB 2; Length 567;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 NSNRAHQKHETGYHKT 17
||||| ||| |||
DB 46 NSNRSKNGKSMFHK 62

RESULT 2
A48462
dense granule protein GRA4 - Toxoplasma gondii
C/Species: Toxoplasma gondii
C/Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
C/Accession: A48462
Mol. Biochem. Parasitol. 56, 227-238, 1992
A/Title: Molecular cloning of GRA4, a Toxoplasma gondii dense granule protein, recogn
A/Reference number: A48462; MIM:3133211; PMID:1362450
A/Accession: A48462
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-345 <MEV>
A/Cross-references: GB:M76432; NID:g161921; PID:g161922
A/Note: sequence extracted from NCBI backbone (NCBI:122806, NCBI:122808)

Query Match 42.0%; Score 47; DB 2; Length 345;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 3; Indels 4; Gaps 1;
OY 6 HGRK---HETGYHKTPT 19
||| | ||:|||

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:42:25 ; Search time 2.74444 Seconds

(without alignments)
287.144 Million cell updates/sec

Title: US-09-834-794A-5
Perfect score: 112
Sequence: 1 NSNRHGGKHEHYGHKTPY 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	112	100.0	127 1 SY28_HUMAN	G9NRJ3 homo sapien
2	47	42.0	345 1 GRA4_TOXO	Q27002 toxoplasma
3	46	42.0	719 1 YP62_CAEEL	Q09437 caenorhabd1
4	46	41.1	260 1 MUR1_CLOPE	Q8XHB7 clostridium
5	44	39.3	630 1 SG64_HUMAN	P31645 homo sapien
6	44	39.3	630 1 SG64_HUMAN	Q9MYX0 macaca mula
7	43.5	38.8	277 1 KNOB_PLAFD	P05229 plasmodium
8	43.5	38.8	473 1 KNOB_PLAFD	P13817 plasmodium
9	43.5	38.8	634 1 KNOB_PLAFD	P09346 plasmodium
10	43.5	38.8	657 1 KNOB_PLAFD	P06719 plasmodium
11	43	38.4	630 1 SG64_MOUSE	O60857 mus muscula
12	43	38.4	757 1 ECR_LUCCU	O18531 lucilia cup
13	43	38.4	991 1 DHP1_SCHPO	P40848 schizosacch
14	42	37.5	336 1 FILA_MOUSE	P11088 mus musculi
15	42	37.5	367 1 VE2_HPV54	O81021 human papil
16	42	37.5	554 1 GYRA_MCFV	Q49166 mycobacteri
17	41.5	37.1	335 1 YG25_YEAST	P53338 saccharomyc
18	41.5	37.1	883 1 DAG1_MOUSE	O62165 mus musculi
19	41.5	37.1	885 1 DAG1_MOUSE	O18738 bos taurus
20	41.5	37.1	895 1 DAG1_HUMAN	O14118 homo sapien
21	41.5	37.1	895 1 DAG1_HUMAN	O28685 erythrocytag
22	41	36.6	96 1 TEN2_TENMO	Q27720 tenebrio mo
23	41	36.6	113 1 YJ52_YEAST	P47128 saccharomyc
24	41	36.6	212 1 MSRA_SALTI	O84150 salmoneilla
25	41	36.6	212 1 MSRA_SALTI	O84171 bacillus su
26	41	36.6	547 1 CATX_HACSU	P09609 kluyveromyc
27	41	36.6	687 1 GALX_KIULA	P70478 rattus norv
28	41	36.6	2842 1 APC_RAT	P21749 drosophila
29	40	35.7	139 1 SALA_DROSI	P21750 drosophila
30	40	35.7	142 1 SALA_DROME	P21748 drosophila
31	40	35.7	142 1 SALA_DROSI	P15834 human immun
32	40	35.7	216 1 VIF_HV2D2	O91155 cyandium c
33	40	35.7	275 1 RK2_CYACA	

34	40	35.7	375 1	LDRL_MOUSE	P70662 mus musculu
35	40	35.7	375 1	LDRL_XENLA	P70060 xenopus lae
36	40	35.7	478 1	RIR3_RAT	O942P5 rattus norv
37	40	35.7	491 1	ZAPA_PROMI	O11137 proteus mir
38	40	35.7	549 1	AAK1_RAT	P54645 rattus norv
39	40	35.7	550 1	DSX_DROME	P23023 rattus norv
40	40	35.7	621 1	CAEL_HUMAN	O13131 homo sapien
41	40	35.7	740 1	YCG2_CAEEL	P51115 caenorhabd1
42	40	35.7	740 1	PSMB_SYNE	P25897 synechococc
43	40	35.7	988 1	SYN3	P73505 synechococc
44	40	35.7	1253 1	DSEP_HUMAN	O92444 homo sapien
45	39.5	35.3	457 1	VIPR_MELGA	O91085 melagris g

ALIGNMENTS

RESULT 1
SY28_HUMAN
ID SY28_HUMAN STANDARD; PRT; 127 AA.

AC Q9NRJ3:
DR 16-OCT-2001 (Rel. 40, Created)
DR 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine A28 precursor (CCL28) (Mucosae-associated epithelial chemokine) (Mec) (CCM1 protein).
CN SCYA28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RC TISSUE=Fetal heart, and Osteoblast.
RX MEDLINE=20357357; PubMed=10781587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Homey B., Catron D., Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J., Kershenovich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A., "Identification of a novel chemokine (CCL28), which binds CCR10 (GPR2)."
RT J. Biol. Chem. 275:22313-22323(2000) /
RN [2]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RX MEDLINE=20432268; PubMed=10975800.
RA Pan J., Kunkel E.J., Gossard D., Lazarus N., Langdon P., Broadwell K., Viera M.A., Genovese M.C., Butcher E.C., Soler D.,
RT "A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial cells in mucosal tissues."
RL J. Immunol. 165:2943-2949(2000) /
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang W., He L., Yuan Z., Wan T., Cao X.,
RT "A novel CC chemokine homology with TECK."
Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
CC - FUNCTION: CHEMOTACTIC ACTIVITY FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS. BINDS TO CCR3 AND CCR10 AND INDUCES CALCIUM MOBILIZATION IN A DOSE-DEPENDENT MANNER.
CC - SUBCELLULAR LOCATION: Secreted.
CC - TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED BY EPITHELIAL CELLS OF DIVERSE TISSUES INCLUDING NORMAL AND PATHOLOGICAL COLON, SALIVARY GLAND, MAMMARY GLAND, TRACHEA AND RECTUM. ALSO FOUND IN PROSTATE, SPLEEN, THYROID, PSORIASIS SKIN AND IN LOWER LEVELS IN PERIPHERAL BLOOD LEUKOCYTES, SMALL INTESTINE, PEYER'S PATCHES, STOMACH AND NORMAL SKIN.
CC - SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE C-C) (CHEMOKINE CC).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-cc the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@lsb-sib.ch).

CC EMBL: AF220210; AAF87205.1; -
 CC EMBL: AF266504; AAG16691.1; -
 DR EMBL: AF110384; AAG41193.1; -
 DR MIM: 605240; -
 DR InterPro: IPR000827; CC.ChemKine.sml.
 DR InterPro: IPR001811; ChemKine_IL8.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
 KW Cytokine; Chemotaxis; Signal; Glycoprotein.
 FT SIGNAL 1 19
 FT CHAIN 20 127
 FT DISULFID 30 58
 FT DISULFID 31 73
 FT CARBOHYD 78 78
 FT SEQUENCE 127 AA; 14280 MW; 3E851A63A2C6D62 CRC64; (POTENTIAL).

Query Match 100.0%; Score 112; DB 1; Length 127;
 Best Local Similarity 100.0%; Pred. No. 7e-11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGKTPY 19
 DB 109 NSNRHOGKHETGKTPY 127

RESULT 2
 GRA4_TOXGO STANDARD; PRT: 345 AA.
 ID Q27002; Q03883; O15889;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Dense granule protein 4 precursor (Protein GRA 4) (Antigen H11).
 GN GRA OR H11.
 OS Toxoplasma gondii.
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida; Sarcocystidae;
 OC Toxoplasma.
 OX NCBI_TaxID=5811;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RH:
 RA MEDLINE-93133211; PubMed-1362450;
 RA Mevelec M.-N., Charles T., Mercereau-Pujalon O., Bourguin I.,
 RA Acharrou A., Dubremetz J.-F., Bout D.;
 RT "Molecular cloning of GRA4, a Toxoplasma gondii dense granule
 RT protein, recognized by mucosal IgA antibodies."
 RT Mol. Biochem. Parasitol. 56:227-238(1992).
 RL [2]
 RP SEQUENCE OF 293-345 FROM N.A.
 RC STRAIN-RH:
 RA MEDLINE-91216437; PubMed-2022319;
 RA Johnson A.M., Illana S.;
 RT "Cloning of Toxoplasma gondii gene fragments encoding diagnostic
 RT antigens."
 RT Gene 99:127-132(1991).
 RL [1]
 RP FUNCTION: MAJOR GRANULAR COMPONENT INVOLVED IN EXCRETED-SECRETED
 CC ANTIGEN (ESA) IMMUNITY.
 CC -1 SUBCELLULAR LOCATION: LOCATED IN DENSE GRANULES OF TACHYZOITES.
 CC UPON INFECTION, SECRETED INTO THE PARASITOPHOROUS VACUOLE (PV) AND
 CC TARGETED TO THE MICROVILLOS MEMBRANOUS NETWORK.
 CC -1 PTM: MAY BE O-GLYCOSYLATED.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

EMBL: M76432; AAA30142.1; -

DR EMBL: M57303; AAA30137.1; ALT-SEQ.
 KW Glycoprotein; Antigen; Transmembrane; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 345
 FT TRANSMEM 276 345
 FT DOMAIN 202 297
 FT DOMAIN 243 248
 FT DOMAIN 287 295
 FT CONFLICT 290 292
 FT SEQUENCE 345 AA; 36283 MW; AC64D1EBE58569 CRC64;

Query Match 42.0%; Score 47; DB 1; Length 345;
 Best Local Similarity 50.0%; Pred. No. 3.5;
 Matches 9; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

OY 6 HOGK----HETGKTPY 19
 DB 44 HQAKRRYHSNMGNOTPY 61

RESULT 3
 YP62_CAEEL STANDARD; PRT: 719 AA.
 ID YP62_CAEEL
 AC Q09437;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative serine/threonine-protein kinase B0495.2 in chromosome II
 DE (EC 2.7.1.-).
 GN B0495.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Kitten J.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).

EMBL: U21317; AAA62523.1; -
 DR HSSP: P24941; 1A01.
 DR WormPep: B0495.2; CE01761.
 DR InterPro: IPR000719; Euk.pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 130 138
 FT DOMAIN 276 279
 FT DOMAIN 356 647
 FT NP_BIND 362 370
 FT BINDING 385 385
 FT ACT_SITE 484 484
 FT SEQUENCE 719 AA; 83567 MW; 219053281FECF15D CRC64;

Query Match 42.0%; Score 47; DB 1; Length 719;
 Best Local Similarity 66.7%; Pred. No. 7.6;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 HOGKHETGHT 17
 DB 137 HRRKHETGHT 148

RESULT 4
 MURI.CLOPE
 ID MURI.CLOPE STANDARD; PRT; 260 AA.
 AC 08XHB7;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Glutamate racemase (EC 5.1.1.3).
 GN MURI OR CPE2568.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium
 NCBI_TaxID=1502;
 RX STRAIN-13 / Type A;
 RA Shintzu T., Ohnami K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh-eater."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 CC -1- FUNCTION: Provides the (R)-glutamic acid required for cell wall
 CC biosynthesis (By similarity).
 CC -1- CATALYTIC ACTIVITY: L-glutamate -> D-glutamate.
 CC -1- PATHWAY: Peptidoglycan biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ASPARTATE/GLUTAMATE RACEMASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AP00194; BAB2274.1;
 CC InterPro: IPR001920; Asp/Glu_race.
 CC DR TIGR00067; glu_race.1.
 CC DR PROSITE: PS00923; ASP_GLU_RACEMASE_1; 1.
 CC DR PROSITE: PS00924; ASP_GLU_RACEMASE_2; 1.
 CC DR Peptidoglycan synthesis; Cell wall; Isomerase; Complete proteome.
 CC SEQUENCE 260 AA; 29177 MW; 2CFCFBB64CFCBC7 CRC64;

Query Match 41.1%; Score 46; DB 1; Length 260;
 Best Local Similarity 50.0%; Pred. No. 3.8;
 Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

OY 1 NSNRANGG-KHETGHT 18
 DB 32 NENYIYGDCKHAPYGEKTP 51

RESULT 5
 S6A4_HUMAN
 ID S6A4_HUMAN STANDARD; PRT; 630 AA.
 AC P31645;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE Sodium-dependent serotonin transporter (5HT transporter) (5HTT).
 GN SLC6A4 OR HTP OR SERT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-9319744; PubMed-8452685;
 RA Lesch K.P., Molozin B.L., Estler H.C., Murphy D.L., Riederer P.;
 RT "Isolation of a cDNA encoding the human brain serotonin transporter."
 RL J. Neural Transm. 91:67-73(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93211998; PubMed-7681602;
 RA Ramamoorthy S., Bauman A.L., Moore K.R., Han H., Yang-Peng T.,
 RA Chang A.S., Ganapathy V., Blakely R.D.;
 RT "Antidepressant- and cocaine-sensitive human serotonin transporter:
 RT molecular cloning, expression, and chromosomal localization."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2542-2546(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Platelet;
 RX MEDLINE-93260476; PubMed-7684072;
 RA Lesch K.P., Molozin B.L., Murphy D.L., Riederer P.;
 RT "Primary structure of the human platelet serotonin uptake site:
 RT identity with the brain serotonin transporter."
 RL J. Neurochem. 60:2319-2322(1993).
 RN [4]
 RP SEQUENCE OF 1-114 FROM N.A.
 RC TISSUE-Placenta;
 RA Bradley C.C., Blakely R.D.;
 RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH
 CC AFFINITY SODIUM-DEPENDENT RECEPTOR INTO PRESYNAPTIC TERMINALS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR
 CC STIMULANTS SUCH AS AMPHETAMINES OR COCAINE.
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
 CC FAMILY (SNF).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X70697; CAAS0029.1;
 CC EMBL: L05568; AA35492.1;
 CC EMBL: U79746; AAB93475.1;
 CC PIR: S37688; S37688.
 CC PIR: A47398; A47398.
 CC DR Genew; HGNC:11050; SLC6A4.
 CC MIM: 182138;
 CC DR InterPro: IPR002437; 5HT_transporter.
 CC DR InterPro: IPR000175; Na/ntiran_symport.
 CC DR Pfam: PF00209; SNF_1.
 CC DR Pfam: PF03491; 5HT_transporter_1.
 CC DR PRINTS: PR00176; NANEUSMPORT.
 CC DR ProDom: PD000448; Na/ntiran_symport_1.
 CC DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
 CC DR PROSITE: PS00754; NA_NEUROTRAN_SYM_2; 1.
 CC DR PROSITE: PS00267; NA_NEUROTRAN_SYM_3; 1.
 CC KW Neurotransmitter transport; Transport; Transmembrane; Glycoprotein;
 CC Symport.
 CC
 CC DOMAIN
 CC FT DOMAIN 1 87 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 88 108 1 (POTENTIAL).
 CC FT TRANSMEM 116 135 2 (POTENTIAL).
 CC FT TRANSMEM 160 180 3 (POTENTIAL).
 CC FT DOMAIN 181 252 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 253 271 4 (POTENTIAL).
 CC FT TRANSMEM 280 297 5 (POTENTIAL).
 CC FT TRANSMEM 333 350 6 (POTENTIAL).
 CC FT TRANSMEM 362 383 7 (POTENTIAL).
 CC FT TRANSMEM 417 436 8 (POTENTIAL).
 CC FT TRANSMEM 464 482 9 (POTENTIAL).

FT TRANSMEM 498 518 10 (POTENTIAL).
 FT TRANSMEM 539 558 11 (POTENTIAL).
 FT TRANSMEM 577 595 12 (POTENTIAL).
 FT DOMAIN 596 630 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 630 AA: 70324 MW: 0EB3550A5799DA2 CRC64;

Query Match 39.3%; Score 44; DB 1; Length 630;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 6 HOGKETHYGHKTPY 19
 DB 75 HOGERTWCKKYDF 88

RESULT 6

66A4_MACMU STANDARD: PRT: 630 AA.

ID 66A4_MACMU

AC 09MYX0:

DT 16-OCT-2001 (Rel. 40; Created)

DT 16-OCT-2001 (Rel. 40; Last sequence update)

DT 15-JUN-2002 (Rel. 41; Last annotation update)

DE Sodium-dependent serotonin transporter (5HT transporter) (5HTT).

GN SLCA4.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Midbrain;

RA MEDLINE=21126513; PubMed=11223167;

RT "Cloning of dopamine, norepinephrine and serotonin transporters from monkey brain: relevance to cocaine sensitivity.";

RL Brain Res. Mol. Brain Res. 87:124-143(2001).

CC -1- FUNCTION: TERMINATES THE ACTION OF SEROTONINE BY ITS HIGH AFFINITY SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS (BY similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- MISCELLANEOUS: THIS PROTEIN IS THE TARGET OF PSYCHOMOTOR STIMULANTS SUCH AS AMPHETAMINES OR COCAINE (BY similarity).

CC -1- SIMILARITY: BELONGS TO THE SODIUM/NEUROTRANSMITTER SYMPORTER FAMILY (SMF).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

FT TRANSMEM 253 271 4 (POTENTIAL).
 FT TRANSMEM 280 297 5 (POTENTIAL).
 FT TRANSMEM 333 350 6 (POTENTIAL).
 FT TRANSMEM 362 383 7 (POTENTIAL).
 FT TRANSMEM 417 436 8 (POTENTIAL).
 FT TRANSMEM 464 482 9 (POTENTIAL).
 FT TRANSMEM 498 518 10 (POTENTIAL).
 FT TRANSMEM 539 558 11 (POTENTIAL).
 FT TRANSMEM 577 595 12 (POTENTIAL).
 FT DOMAIN 596 630 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 630 AA: 70257 MW: 79DFB7E5E9B692A CRC64;

Query Match 39.3%; Score 44; DB 1; Length 630;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 6 HOGKETHYGHKTPY 19
 DB 75 HOGERTWCKKYDF 88

RESULT 7

66A4_PLAFD STANDARD: PRT: 277 AA.

ID 66A4_PLAFD

AC P05229;

DT 13-AUG-1987 (Rel. 05; Created)

DT 13-AUG-1987 (Rel. 05; Last sequence update)

DT 01-OCT-1996 (Rel. 34; Last annotation update)

DE Plasmidum falciparum histidine-rich protein (KAHP) (Fragment).

OS Plasmidum falciparum (Isolate CDC / Honduras).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmidium.

OX NCBI_TaxID=5836;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=87275837; PubMed=3301326;

RA Ardeshtir F., Flint J.E., Matsumoto Y., Akawa M., Reese R.T., Stanley H.;

RT "cdna sequence encoding a plasmidum falciparum protein associated with knobs and localization of the protein to electron-dense regions in membranes of infected erythrocytes.";

RL EMBO J. 6:1421-1427(1987).

CC -1- FUNCTION: KAHP MIGHT MIMICK HUMAN HISTIDINE-RICH GLYCOPROTEINS TO ANCHOR HOST THROMBOSPONDIN OR A PARASITE ANALOG IN A BINDING COMPLEX WITH THE ENDOHELIAL CELL RECEPTOR.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF THE MEMBRANE OF INFECTED ERYTHROCYTES.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

```

RESULT 8
ID KNOB_PLAFA STANDARD; PRT: 473 AA.
AC P13817;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE knob-associated histidine-rich protein precursor (KAHRP) (HRPI)
DE (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8812262; PubMed=2448620;
RA Ellis J., Irving D.O., Wellens T.E., Howard R.J., Cross G.A.M.;
RT "Structure of the knob-associated histidine-rich
RT protein of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 26:203-214(1987).
CC -1- FUNCTION: KAHRP MIMICK HUMAN HISTIDINE-RICH GLYCOPROTEINS
CC TO ANCHOR HOST THROMBOSPONDIN OR A PARASITE ANALOG IN A BINDING
CC COMPLEX WITH THE ENDOTHELIAL CELL RECEPTOR.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF THE MEMBRANE OF INFECTED
CC ERYTHROCYTES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: M19028; AAA29630.1; -.
DR PIR: A54494; A54494.
KW Malaria; Repeat; Signal.
FT SIGNAL 1 34
FT CHAIN 35 >473 POTENTIAL.
FT DOMAIN 61 116 KNOB-ASSOCIATED HISTIDINE-RICH PROTEIN.
FT CARBOHYD 42 42 HIS-RICH.
FT NON_TER 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 473 AA; 53191 MW; A97F2D3DA7675E2E CRC64;
SQ
Query Match 38.8%; Score 43.5; DB 1; Length 473;
Best Local Similarity 34.6%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 4; Indels 11; Gaps 1;
OY 5 AHQGR-----ETYGKTRPY 19
DB 278 AHDSHGRLRGHGKSGEGYEAPY 303
RESULT 9
ID KNOB_PLAFA STANDARD; PRT: 634 AA.
AC P09346; P13818;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE knob-associated histidine-rich protein precursor (KAHRP) (KP).
OS Plasmodium falciparum (isolate FCR-3 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5838;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88041063; PubMed=3313387;
RA Pologe L.G., Pavlovic A., Shio H., Ravetch J.V.;
RT "Primary structure and subcellular localization of the
RT knob-associated histidine-rich protein of Plasmodium falciparum."
RL Proc. Natl. Acad. Sci. U.S.A. 84:7139-7143(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88122251; PubMed=2828945;

```

```

RA Sharma Y.D., Kilejian A.;
RT "Structure of the knob protein (KP) gene of Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 26:11-16(1987).
RN [3]
RP SEQUENCE OF 1-270 FROM N.A.
RX MEDLINE=87017062; PubMed=3532126;
RA Kilejian A., Sharma Y.D., Karouli H., Naslund L.;
RT "Histidine-rich domain of the knob protein of the human malaria
RT parasite Plasmodium falciparum."
RL Proc. Natl. Acad. Sci. U.S.A. 83:7938-7941(1986).
CC -1- FUNCTION: KAHRP MIMICK HUMAN HISTIDINE-RICH GLYCOPROTEINS
CC TO ANCHOR HOST THROMBOSPONDIN OR A PARASITE ANALOG IN A BINDING
CC COMPLEX WITH THE ENDOTHELIAL CELL RECEPTOR.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SIDE OF THE MEMBRANE OF INFECTED
CC ERYTHROCYTES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: J02972; AAA29629.1; -.
DR EMBL: M19881; AAA29632.1; -.
DR PIR: M14210; AAA29631.1; -.
DR PIR: A28412; A28412.
DR PIR: A26480; A26480.
DR PIR: A54495; A54495.
KW Malaria; Repeat; Signal.
FT SIGNAL 1 34
FT CHAIN 35 634
FT DOMAIN 61 116 HIS-RICH.
FT DOMAIN 370 438 13 X 10 AA TANDEM REPEATS.
FT DOMAIN 540 580 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 42 42 V -> T (IN REF. 2 AND 3).
FT CONFLICT 260 260 A -> T (IN REF. 2).
FT CONFLICT 492 492 V -> A (IN REF. 2).
FT CONFLICT 507 507 P -> T (IN REF. 2).
FT CONFLICT 539 539 A -> T (IN REF. 2).
FT CONFLICT 543 543 A -> T (IN REF. 2).
FT SEQUENCE 634 AA; 69149 MW; 4D5E28EA6171DE7D CRC64;
SQ
Query Match 38.8%; Score 43.5; DB 1; Length 634;
Best Local Similarity 34.6%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 4; Indels 11; Gaps 1;
OY 5 AHQGR-----ETYGKTRPY 19
DB 278 AHDSHGRLRGHGKSGEGYEAPY 303
RESULT 10
ID KNOB_PLAFA STANDARD; PRT: 657 AA.
AC P06719;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE knob-associated histidine-rich protein precursor (KAHRP).
OS Plasmodium falciparum (isolate NF7 / Ghana).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87275836; PubMed=3301325;
RA Trigila T., Stahl H.-D., Creweher P.E., Scanlon D.B., Brown G.V.,
RA Anders R.F., Kemp D.J.;
RT "The complete sequence of the gene for the knob-associated histidine-
RT rich protein from Plasmodium falciparum."
RL EMBO J. 6:1413-1419(1987).

```


DB 75 HOGRETNWKKMDF 88

RESULT 12

ECR_LUCCU

ID ECR_LUCCU STANDARD; PRT; 757 AA.

AC 018531;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ecdysone receptor (Ecdysteroid receptor) (20-hydroxy-ecdysone receptor) (20E receptor).

GN ECR OR NR1H1.

OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea; Calliphoridae; Lucilia.

OX NCBI_TaxID=7375;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97449774; PubMed=9304790;

RA Hannan G.N., Hill R.J.;

RT Cloning and characterization of LcECR: a functional ecdysone receptor from the sheep blowfly *Lucilia cuprina*.;

RL Insect Biochem. Mol. Biol. 27:479-488(1997).

CC -1- FUNCTION: RECEPTOR FOR ECDYSONE. BINDS TO ECDYSONE RESPONSE ELEMENTS (ECRES) (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC NRI SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U75355; AAB81130.1; -.

DR HSSP: P20393; 1A6Y.

DR InterPro: IPR000536; Hormone_rec_1lg.

DR InterPro: IPR001723; Stdhnm_receptor.

DR InterPro: IPR001628; znf_C4steroid.

DR Pfam: PF00104; hormone_rec; 1.

DR Pfam: PF00105; zf-C4; 1.

DR PRINTS: PR00398; STRDHOMONER.

DR PRINTS: PR00047; STRDHOFINGER.

DR ProDom: PD000035; znf_C4steroid; 1.

DR SMART: SM00399; znf_C4; 1.

DR SMART: SM00430; HOI1; 1.

DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein; zinc-finger.

KM

FT DOMAIN 1 300 MODULATING (POTENTIAL).

FT DNA_BIND 301 366 NUCLEAR RECEPTOR-TYPE.

FT ZN_FING 301 321 C4-TYPE.

FT ZN_FING 337 361 C4-TYPE.

FT DOMAIN 454 674 HORMONE-BINDING (POTENTIAL).

FT SEQUENCE 757 AA; 83075 MW; C1511452ED37D359 CRC64;

SO

Query Match 38.4%; Score 43; DB 1; Length 757;

Best Local Similarity 46.7%; Pred. No. 34;

Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 NSNRHOGKRETYGH 15

DB 169 NSNHSNSSSHRTNGH 183

RESULT 13

DHP1_SCHPO STANDARD; PRT; 991 AA.

ID DHP1_SCHPO

P40848;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Protein dhp1.

GN DHP1 OR SPAC26A3.12.

OS Schizosaccharomyces pombe (fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-29.

RC STRAIN=975;

RX MEDLINE=94247347; PubMed=8190062;

RA Sugano S., Shobike T., Takeda T., Sugino A., Ikeda H.;

RT "Molecular analysis of the dhp1+ gene of *Schizosaccharomyces pombe*: an essential gene that has homology to the DSTR and RAT1 genes of *Saccharomyces cerevisiae*.";

RT Mol. Gen. Genet. 243:1-8(1994).

RL

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris N., Huckle E.J., Hunt S., Jagels K., Holtroyd S., Hornsby T., Howarth S., Mungall K., Mungall L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Shelson J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Voicikart G., Aert R., Robben J., Gymnopoulos B., Wellitsch I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H., Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M., Beger P., Zimmermann W., Wedler H., Wambut R., Purnelle B., Goftau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottler S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hunt S.M., Lucas M., Rochet M., Galliaridin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shipakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of *Schizosaccharomyces pombe*.";

RL Nature 415:871-880(2002).

RN [3]

RP CHARACTERIZATION.

RX MEDLINE=21138436; PubMed=11238999;

RA Shobike T., Takedayashi K., Tan T., Sugano S., Ikeda H.;

RT "The dhp1+ gene, encoding a putative nuclear 5'3' exonuclease, is required for proper chromosome segregation in fission yeast.";

RL Nucleic Acids Res. 29:1326-1333(2001).

CC -1- FUNCTION: ESSENTIAL FOR VEGETATIVE CELL GROWTH; REQUIRED FOR PROPER CHROMOSOME SEGREGATION. CAN BIND TO DNA AND HAS 5'->3' EXORIBONUCLEASE ACTIVITY.

CC -1- SUBUNIT: Interacts with dnl1.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- SIMILARITY: TO YEAST RAT1; SOME, TO YEAST KEM1/DST2 AND S.POMBE EXO2.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: D17752; BAA04601.1; -
 DR EMBL: Z69240; CAA93235.1; -
 DR PIR: S43891; S43891.
 DR InterPro: IPR004859; Put_53exo.
 DR Pfam: PF03159; DUF251; 1.
 KM Nucleic protein: Hydrolyase; Nuclease; Exonuclease.
 FT DOMAIN 264 268
 SQ SEQUENCE 991 AA; 112367 MW; 83FA34D93DADHD00 CRC64;

Query Match 38.4%; Score 43; DB 1; Length 991;
 Best Local Similarity 31.6%; Pred. No. 45;
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 NSNRHOGKHETYGKTPY 19
 DB 881 NNGNSYRGHQSXYGRSY 899

RESULT 14

FILE_MOUSE STANDARD; PRT; 336 AA.

AC P1088;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Flaggrin (Fragment).
 GN FLG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:88058903; PubMed-3680218;
 RA Rothnagel J.A., Mehrel T., Idler W.W., Roop D.R., Steinert P.M.;
 RT "The gene for mouse epidermal flaggrin precursor. Its partial
 RT characterization, expression, and sequence of a repeating flaggrin
 RT unit."
 RL J. Biol. Chem. 262:15643-15648(1987).
 RN [2]
 RP REVISIONS.

RA Rothnagel J.A.;
 RL Submitted (SEP-1988) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: AGGREGATES KERATIN INTERMEDIATE FILAMENTS AND PROMOTES
 CC DISULFIDE-BOND FORMATION AMONG THE INTERMEDIATE FILAMENTS DURING
 CC TERMINAL DIFFERENTIATION OF MAMMALIAN EPIDERMIS.
 CC -1- PFM: FLAGGRIN IS INITIALLY SYNTHESIZED AS A LARGE, INSOLUBLE,
 CC HIGHLY PHOSPHORYLATED PRECURSOR CONTAINING MANY TANDEM COPIES
 CC OF 248 AA, WHICH ARE NOT SEPARATED BY "LARGE LINKER". THE
 CC PRECURSOR IS DEPOSITED AS KERATOHALIN GRANULES. DURING TERMINAL
 CC DIFFERENTIATION IT IS DEPHOSPHORYLATED & PROTEOLYTICALLY CLEAVED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: J03458; AAA75559.1; -
 DR PIR: A28444; A28444.
 DR MGD: MGI:95553; Flg.
 KM Phosphorylation; Developmental protein.
 FT NON_TER 1
 SQ SEQUENCE 336 AA; 35678 MW; 259F124D3AC0DB2D CRC64;

Query Match 37.5%; Score 42; DB 1; Length 336;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 NSNRHOGKHETYG 14
 : | | | | | | |

DB 200 SSGRGHGAHQEOG 213

RESULT 15

VE2_HPVS4 STANDARD; PRT; 367 AA.

AC Q81021;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Regulatory protein E2.

GN E2.
 OS Human papillomavirus type 54.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID-37113;
 RN [1]

RP SEQUENCE FROM N.A.

RA delius H.;
 RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.

CC -1- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGCT-3') PRESENT
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDENT OF E2RE'S POSITION
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA
 CC REPLICATION.
 CC -1- SUBUNIT: BINDS DNA AS A DIMER.
 CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U37488; AAA79190.1; -
 DR HSP: P17383; IDHM.
 DR InterPro: IPR000427; E2_C.
 DR InterPro: IPR001866; E2_N.
 DR Pfam: PF00508; E2_N; 1.
 DR Pfam: PF00511; E2_C; 1.
 DR Prodom: PD00672; E2_C; 1.
 DR Prodom: PD00678; E2_N; 1.
 KM Early protein; Transcription regulation; Activator; DNA-binding;
 KM Trans-acting factor; DNA replication; Repressor; Nuclear protein.
 SQ SEQUENCE 367 AA; 41690 MW; 77FB3CF03EDD674 CRC64;

Query Match 37.5%; Score 42; DB 1; Length 367;
 Best Local Similarity 38.9%; Pred. No. 23;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 NSNRHOGKHETYGKTP 18
 DB 271 DDNNKNGRHTSSCDTTP 288

Search completed: January 14, 2003, 18:13:52
 Job time : 3.74444 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:15:35 ; Search time 2.53333 Seconds
(without alignments)
145.508 Million cell updates/sec

Title: US-09-834-794a-5
Perfect score: 112
Sequence: 1 NSNRHOGKHETYGKTPY 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues
Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, AA:*

- 1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubppa/PCP_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubppa/PCPUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	19	10 US-09-834-795a-5	Sequence 5, Appl1
2	112	100.0	104	10 US-09-834-795a-2	Sequence 2, Appl1
3	112	100.0	127	9 US-09-898-751a-6	Sequence 6, Appl1
4	112	100.0	127	10 US-09-813-492-2	Sequence 2, Appl1
5	112	100.0	127	10 US-09-834-795a-1	Sequence 1, Appl1
6	112	100.0	127	10 US-09-831-381a-2	Sequence 2, Appl1
7	44	39.3	630	10 US-09-843-598-10	Sequence 10, Appl1
8	42.5	37.9	325	10 US-09-864-761-36983	Sequence 36983, A
9	42	37.5	129	10 US-09-815-242-13322	Sequence 13322, A
10	42	37.5	129	10 US-09-815-242-13576	Sequence 13576, A
11	42	37.5	892	9 US-09-895-913a-198	Sequence 198, App
12	41	36.6	431	12 US-10-062-254-272	Sequence 272, App
13	41	36.6	832	10 US-09-919-585-21	Sequence 20, Appl1
14	40	35.7	426	9 US-09-860-846-20	Sequence 20, Appl1
15	40	35.7	426	10 US-09-861-289-20	Sequence 20, Appl1
16	40	35.7	466	10 US-09-925-301-1366	Sequence 1366, Ap
17	40	35.7	796	10 US-09-205-658-40	Sequence 40, Appl1
18	40	35.7	796	10 US-09-844-353a-40	Sequence 40, Appl1
19	40	35.7	858	10 US-09-205-658-41	Sequence 41, Appl1

20	40	35.7	858	10 US-09-844-353a-41	Sequence 41, Appl1
21	40	35.7	892	10 US-09-205-658-42	Sequence 42, Appl1
22	40	35.7	892	10 US-09-844-353a-42	Sequence 42, Appl1
23	40	35.7	950	10 US-09-823-356-9	Sequence 9, Appl1
24	40	35.7	3782	9 US-09-860-846-4	Sequence 4, Appl1
25	40	35.7	3782	10 US-09-861-289-4	Sequence 4, Appl1
26	39.5	35.3	344	10 US-09-815-242-5059	Sequence 5059, Ap
27	39	34.8	767	9 US-09-987-482-3	Sequence 3, Appl1
28	39	34.8	842	10 US-09-798-831-8	Sequence 8, Appl1
29	39	34.8	1643	10 US-09-515-806-2	Sequence 2, Appl1
30	39	34.8	2843	8 US-08-681-219-32	Sequence 32, Appl1
31	39	34.8	2843	9 US-09-987-482-1	Sequence 21, Appl1
32	38.5	34.4	180	10 US-09-997-701-5	Sequence 5, Appl1
33	38.5	34.4	247	10 US-09-452-239-18	Sequence 18, Appl1
34	38	33.9	71	9 US-10-001-835-211	Sequence 211, App
35	38	33.9	199	9 US-09-764-868-1137	Sequence 1137, Ap
36	38	33.9	551	10 US-09-818-143-21	Sequence 21, Appl1
37	38	33.9	616	9 US-09-738-626-5189	Sequence 5189, Ap
38	38	33.9	2175	10 US-09-935-541-2	Sequence 2, Appl1
39	38	33.9	2188	10 US-09-935-541-4	Sequence 4, Appl1
40	37.5	33.5	520	9 US-09-738-626-5981	Sequence 5981, Ap
41	37.5	33.5	1687	9 US-10-094-679-3	Sequence 3, Appl1
42	37.5	33.5	2597	10 US-09-905-129-2	Sequence 2, Appl1
43	37.5	33.5	2597	10 US-09-905-129-10	Sequence 10, Appl1
44	37.5	33.5	2597	10 US-09-905-129-13	Sequence 13, Appl1
45	37.5	33.5	2597	10 US-09-991-630-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-834-795a-5
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 19
TYPE: PRT
ORGANISM: Homo sapiens
US-09-834-795a-5
Query Match 100.0%; Score 112; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NSNRHOGKHETYGKTPY 19
Db 1 NSNRHOGKHETYGKTPY 19
RESULT 2
US-09-834-795a-2
Sequence 2, Application US/09834795A
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster

APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (68)..(68)
OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
US-09-834-795A-2

Query Match 100.0%; Score 112; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSNRHQGHKHYGHTPY 19
|||||

Db 86 NSNRHQGHKHYGHTPY 104
|||||

RESULT 3
US-09-898-751A-6
Sequence 6, Application US/09898751A
Patent No. US20020160024A1
GENERAL INFORMATION:
APPLICANT: Oldham, Elizabeth R.
APPLICANT: Soto, Hortensia
APPLICANT: Liu, Ying
APPLICANT: Hudak, Susan A.
APPLICANT: Homey, Bernhard
APPLICANT: Morales, Janine M.
APPLICANT: Kellerman, Sirid-Almee
APPLICANT: McEvoy, Leslie M.
APPLICANT: Bowman, Edward P.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
FILE REFERENCE: DX0882XK
CURRENT APPLICATION NUMBER: US/09/898,751A
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: US09/471,549
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: US60/136,570
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US60/113,858
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (529)..(529)
OTHER INFORMATION: unknown amino; may be "A", "C", or "G"
US-09-898-751A-6

Query Match 100.0%; Score 112; DB 9; Length 127;

Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NSNRHQGHKHYGHTPY 19
|||||

Db 109 NSNRHQGHKHYGHTPY 127
|||||

RESULT 4
US-09-813-492-2
Sequence 2, Application US/09813492
Patent No. US20020009735A1
GENERAL INFORMATION:
APPLICANT: Labow, Mark A.
APPLICANT: Mickanin, Craig Stephen
APPLICANT: Bhatia, Umesh
TITLE OF INVENTION: MAMMARY GLAND CHEMOKINE
FILE REFERENCE: 12345
CURRENT APPLICATION NUMBER: US/09/813,492
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 127
TYPE: PRT
ORGANISM: HUMAN
US-09-813-492-2

Query Match 100.0%; Score 112; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NSNRHQGHKHYGHTPY 19
|||||

Db 109 NSNRHQGHKHYGHTPY 127
|||||

RESULT 5
US-09-834-795A-1
Sequence 1, Application US/09834795A
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (91)..(91)
OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-834-795A-1

Query Match 100.0%; Score 112; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGKTPY 19
Db 109 NSNRHOGKHETGKTPY 127

RESULT 6
US-09-931-381A-2
Sequence 2, Application US/09931381A
Patent No. US20020137107A1
GENERAL INFORMATION:
APPLICANT: Butcher, Eugene C.
APPLICANT: Kunkel, Eric J.
APPLICANT: Pan, Junliang
APPLICANT: Soler-Ferran, Dulce
TITLE OF INVENTION: Method for Identifying Agents Which
Modulate Chemokine "Mec"-Induced Functions of CCR3 and/or
TITLE OF INVENTION: CCR3
FILE REFERENCE: 1855, 2010-003
CURRENT APPLICATION NUMBER: US/09/931, 381A
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: U.S. 09/638, 914
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(22)
US-09-931-381A-2

Query Match 100.0%; Score 112; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGKTPY 19
Db 109 NSNRHOGKHETGKTPY 127

RESULT 7
US-09-843-598-10
Sequence 10, Application US/09843598
Patent No. US20020010944A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Ranganathan, Rajesh
TITLE OF INVENTION: CESERT GENES, PROTEINS, AND MODULATORY
FILE REFERENCE: 01997/525002
CURRENT APPLICATION NUMBER: US/09/843, 598
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: US 60/200, 549
PRIOR FILING DATE: 2000-04-26
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 630
TYPE: PRT
ORGANISM: Homo sapiens
US-09-843-598-10

Query Match 39.3%; Score 44; DB 10; Length 630;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 6 HOGKHETGKTPY 19
Db 75 HOGKHETGKTPY 88

RESULT 8
US-09-864-761-36983
Sequence 36983, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeomica-x-1
CURRENT APPLICATION NUMBER: US/09/864, 761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180, 312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207, 456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632, 366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263, 6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236, 359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234, 687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608, 408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774, 203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36983
LENGTH: 325
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004129.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: EST HUMAN HIT: BF240303.1, EVALUE 2.00e-96
OTHER INFORMATION: SWISSPROT HIT: P10659, EVALUE 3.80e+00
US-09-864-761-36983

Query Match 37.9%; Score 42.5; DB 10; Length 325;

Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
OY 6 HOG-KHETYGHKTPY 19
| | | | | | | | | |
Db 308 HSGMKHQFYGHNSY 322

RESULT 9
US-09-815-242-13322
Sequence 13322, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13322
LENGTH: 129
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13322
Query Match 37.5%; Score 42; DB 10; Length 129;
Best Local Similarity 46.2%; Pred. No. 8.5;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 7 OCKHETYGHKTPY 19
| | | | | | | | | |
Db 107 KSHRKQGHROPY 119

RESULT 10
US-09-815-242-13576
Sequence 13576, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13576
LENGTH: 129
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13576
Query Match 37.5%; Score 42; DB 10; Length 129;
Best Local Similarity 46.2%; Pred. No. 8.5;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 7 OCKHETYGHKTPY 19
| | | | | | | | | |
Db 107 KSHRKQGHROPY 119

RESULT 11
US-09-895-913A-198
Sequence 198, Application US/09895913A
Patent No. US20020160456A1
GENERAL INFORMATION:
APPLICANT: Kleantous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles
APPLICANT: Tomb, Jean Francois
TITLE OF INVENTION: Identification of Polynucleotides
FILE REFERENCE: Encoding No. US20020160456A1 Helicobacter Polypeptides in t
CURRENT APPLICATION NUMBER: US/09/895,913A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 08/881,227
PRIOR FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 368
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 198
LENGTH: 892
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-895-913A-198
Query Match 37.5%; Score 42; DB 9; Length 892;
Best Local Similarity 63.6%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 8 GKHEGYGHKTPY 18
| | | | | | | | | |
Db 74 GKRLTYHTHP 84

RESULT 12
US-10-063-254-272
Sequence 272, Application US/10062254
Patent No. US20020138882A1
GENERAL INFORMATION:

APPLICANT: Cahoon, Edgar B
APPLICANT: Falco, Rebecca E
APPLICANT: Fang, Yiwen
APPLICANT: Hantke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062,254
SEQUENCE FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/146511
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO 272
LENGTH: 431
TYPE: PRT
ORGANISM: Glycine max
US-10-062-254-272

Query Match 36.6%; Score 41; DB 12; Length 431;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETYG 14
DB 394 DSNCGHYGCHGTC 407

RESULT 13
US-09-919-585-21
Sequence 21, Application US/09919585
Patent No. US20020115167A1
GENERAL INFORMATION:
APPLICANT: Sun, Tian-Qiang
APPLICANT: Feng, Jia-Jia
APPLICANT: Reinhard, Christoph
APPLICANT: Fanel, Wendy J.
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: ISOLATION OF DROSOPHILA AND HUMAN POLYNUCLEOTIDES ENCODING PAR-1
TITLE OF INVENTION: POLYPEPTIDES ENCODED BY THE POLYNUCLEOTIDES AND METHODS UTILIZING
FILE REFERENCE: PP-016093.002/200130.525
CURRENT APPLICATION NUMBER: US/09/919,585
SEQUENCE FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 22

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 832
TYPE: PRT
ORGANISM: Drosophila sp.
US-09-919-585-21

Query Match 36.6%; Score 41; DB 10; Length 832;
Best Local Similarity 40.0%; Pred. No. 82;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETYG 15
DB 150 NFNHSHQHHH 164

RESULT 14
US-09-860-846-20
Sequence 20, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/860,846
SEQUENCE FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 20
LENGTH: 426
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-860-846-20

Query Match 35.7%; Score 40; DB 9; Length 426;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 5 AHOGKHETYGKHTP 18
DB 390 AHRLEETFFGDPPT 403

RESULT 15
US-09-861-289-20
Sequence 20, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438US1
CURRENT APPLICATION NUMBER: US/09/861,289
SEQUENCE FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 20
LENGTH: 426
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-861-289-20

Query Match 35.7%; Score 40; DB 10; Length 426;
Best Local Similarity 50.0%; Pred. No. 59;

Matches	7;	Conservative	3;	Mismatches	4;	Indels	0;	Gaps	0;
Oy	5	AHOGKHETYGHKTP	18						
		: : :							
Db	390	AHRLRETRFGDPTP	403						

Search completed: January 14, 2003, 18:27:30
Job time : 3.53333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:11:25 ; Search time 4.22222 Seconds
(without alignments)
132.403 Million cell updates/sec

Title: US-09-834-794A-5
Perfect score: 112
Sequence: 1 NSNRAHQKHETYGKHTPY 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/p/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/p/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/p/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/p/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/p/ptodata/1/1aa/PCtUS.COMB.pep.*
6: /cgn2_6/p/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	19	4	US-09-146-580-5
2	112	100.0	104	4	US-09-146-580-2
3	112	100.0	127	4	US-09-146-580-1
4	45	40.2	500	1	US-08-117-083-70
5	44	39.3	630	1	US-07-959-943-11
6	41.5	37.1	845	5	PCT-US94-05387-7
7	41.5	37.1	895	1	US-08-123-161A-8
8	41.5	37.1	895	1	US-08-463-278-8
9	41	36.6	707	4	US-09-134-001C-2962
10	40	35.7	257	3	US-09-101-146-6
11	40	35.7	345	3	US-09-101-146-1
12	40	35.7	426	3	US-09-320-878-8
13	40	35.7	426	4	US-09-105-537-20
14	40	35.7	729	4	US-09-625-188-20
15	40	35.7	796	4	US-08-857-076-40
16	40	35.7	858	4	US-08-857-076-41
17	40	35.7	892	4	US-08-857-076-42
18	40	35.7	3782	4	US-09-105-537-4
19	40	34.8	212	1	US-08-915-003-9
20	39	34.8	212	2	US-08-642-247-9
21	39	34.8	370	4	US-09-134-001C-4166
22	39	34.8	1045	1	US-08-452-083-2
23	39	34.8	2311	4	US-08-934-386-9
24	39	34.8	2555	3	US-09-058-489-36
25	39	34.8	2842	1	US-07-741-940-7
26	39	34.8	2842	1	US-08-289-548A-7
27	39	34.8	2842	1	US-08-452-654-7

28	39	34.8	2842	4	US-08-449-731-7	Sequence 7, Appli
29	39	34.8	2843	1	US-07-741-940-2	Sequence 2, Appli
30	39	34.8	2843	1	US-08-289-548A-2	Sequence 2, Appli
31	39	34.8	2843	1	US-08-452-654-2	Sequence 2, Appli
32	39	34.8	2843	1	US-08-452-655B-2	Sequence 2, Appli
33	39	34.8	2843	1	US-08-452-655B-7	Sequence 7, Appli
34	39	34.8	2843	2	US-08-370-235A-2	Sequence 2, Appli
35	39	34.8	2843	3	US-08-450-582-2	Sequence 2, Appli
36	39	34.8	2843	4	US-08-449-731-2	Sequence 2, Appli
37	39	34.8	2843	4	US-08-821-355A-7	Sequence 7, Appli
38	39	34.8	2973	2	US-09-003-687A-7	Sequence 7, Appli
39	39	34.8	2973	4	US-09-136-605-7	Sequence 7, Appli
40	39	34.8	2973	4	US-09-187-331-5	Sequence 5, Appli
41	38.5	34.4	180	3	US-09-470-946-5	Sequence 5, Appli
42	38.5	34.4	180	4	US-09-452-239-18	Sequence 18, Appli
43	38.5	34.4	247	1	US-08-482-282B-6	Sequence 6, Appli
44	38	33.9	72	1	US-08-486-036A-6	Sequence 6, Appli
45	38	33.9	72	1	US-08-486-036A-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT 1
US-09-146-580-5
: Sequence 5, Application US/09146580A
: Patent No. 6306653
: ORGANISM: Homo sapiens
: APPLICANT: Papsidero, Lawrence D
: APPLICANT: Dyster, Lyn M
: TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
: FILE REFERENCE: 200755/1002
: CURRENT APPLICATION NUMBER: US/09/146,580A
: CURRENT FILING DATE: 1998-09-03
: EARLIER APPLICATION NUMBER: 60/071,889
: EARLIER FILING DATE: 1998-01-20
: EARLIER APPLICATION NUMBER: 60/092,155
: EARLIER FILING DATE: 1998-07-09
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 5
: LENGTH: 19
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-146-580-5

Query Match      100.0%; Score 112; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.6e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 NSNRAHQKHETYGKHTPY 19
Db      1 NSNRAHQKHETYGKHTPY 19

RESULT 2
US-09-146-580-2
: Sequence 2, Application US/09146580A
: Patent No. 6306653
: ORGANISM: Homo sapiens
: APPLICANT: Papsidero, Lawrence D
: APPLICANT: Dyster, Lyn M
: TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
: FILE REFERENCE: 200755/1002
: CURRENT APPLICATION NUMBER: US/09/146,580A
: CURRENT FILING DATE: 1998-09-03
: EARLIER APPLICATION NUMBER: 60/071,889
: EARLIER FILING DATE: 1998-01-20
: EARLIER APPLICATION NUMBER: 60/092,155
: EARLIER FILING DATE: 1998-07-09
: NUMBER OF SEQ ID NOS: 18
```

```
SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)
; OTHER INFORMATION: xaa at position 47 is either Arg or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: xaa at position 68 is either Lys or Asn
; US-09-146-580-2

Query Match
Best Local Similarity 100.0%; Score 112; DB 4; Length 104;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSNRHOGKHETYGHTPY 19
DB 86 NSNRHOGKHETYGHTPY 104

RESULT 3
; US-09-146-580-1
; Sequence 1, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; EARLIER FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentlin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 127
; TYPE: PR1
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (70)
; OTHER INFORMATION: xaa at position 70 is either Arg or Gly
; NAME/KEY: UNSURE
; LOCATION: (91)
; OTHER INFORMATION: xaa at position 91 is either Lys or Asn
; US-09-146-580-1

Query Match
Best Local Similarity 100.0%; Score 112; DB 4; Length 127;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSNRHOGKHETYGHTPY 19
DB 109 NSNRHOGKHETYGHTPY 127

RESULT 4
; US-08-117-083-70
; Sequence 70, Application US/08117083
; Patent No. 5719054
; GENERAL INFORMATION:
; APPLICANT: Bouranell, Michael E.
; APPLICANT: Ingalls, Stephen C.
; APPLICANT: Munro, Alan J.
```

```
TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
; TITLE OF INVENTION: Papilloma Virus Proteins
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Walter H. Dreger
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentlin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,083
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-58783
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..500
; OTHER INFORMATION:
; OTHER INFORMATION: /note="xaa refers to stop codon in
; US-08-117-083-70

Query Match
Best Local Similarity 40.2%; Score 45; DB 1; Length 500;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 10 HETYGHTPY 19
DB 16 HETFGNTPY 25

RESULT 5
; US-07-959-943-11
; Sequence 11, Application US/07959943
; Patent No. 5418162
; GENERAL INFORMATION:
; APPLICANT: Blakely, Randy D.
; APPLICANT: Freneau Jr., Robert T.
; APPLICANT: Caron, Marc G.
; TITLE OF INVENTION: Serotonin Transporter CDNA
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth D. Sibley; Bell, Seltzer, Park
; ADDRESS: and
; ADDRESSEE: Gibson
; STREET: Post Office Drawer 31107
; CITY: Raleigh
; STATE: No. 5418162th Carolina
; COUNTRY: U.S.A.
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```


SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/959,943
FILING DATE: 19921014
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5405,38a
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-881-3140
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 630 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-959-943-11

Query Match 39.3%; Score 44; DB 1; Length 630;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 HOGKHETGKTPY 19
DB 75 HOGKERTGKRVDF 88

RESULT 6
PCT-US94-05387-7
Sequence 7, Application PC/TUS9405387
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: AGRIN RECEPTOR
NUMBER OF SEQUENCES: 7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05387
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/149,966
FILING DATE: 10-NOV-1993
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..445
OTHER INFORMATION: /label= HUMDAGI
FEATURE:
OTHER INFORMATION: /note= "human dystroglycan precursor 451-895"
NAME/KEY: Region
LOCATION: 90..109
OTHER INFORMATION: /note= "190 kDa fragment homology"
FEATURE:
NAME/KEY: Region
LOCATION: 162..201
OTHER INFORMATION: /note= "190 kDa fragment homology"
FEATURE:
NAME/KEY: Region
LOCATION: 123..142
OTHER INFORMATION: /note= "190 kDa fragment homology"
FEATURE:

NAME/KEY: Region
LOCATION: 204..219
OTHER INFORMATION: /note= "50 kDa fragment homology"
FEATURE:
NAME/KEY: Region
LOCATION: 226..243
OTHER INFORMATION: /note= "50 kDa fragment homology"
FEATURE:
NAME/KEY: Region
LOCATION: 301..324
OTHER INFORMATION: /note= "Transmembrane"
FEATURE:
NAME/KEY: Region
LOCATION: 345..369
OTHER INFORMATION: /note= "50 kDa fragment homology"
PCT-US94-05387-7

Query Match 37.1%; Score 41.5; DB 5; Length 445;
Best Local Similarity 38.5%; Pred. No. 64;
Matches 10; Conservative 2; Mismatches 5; Indels 9; Gaps 1;

QY 1 NSN-----RAHOGKHETGKHT 17
DB 105 NSNSQMTGCLPDSHVGHKHEFMHAT 130

RESULT 7
US-08-123-161A-8
Sequence 8, Application US/08123161A
Patent No. 5449616
GENERAL INFORMATION:
APPLICANT: Campbell, Kevin P.
APPLICANT: Roberts, Steven L.
APPLICANT: Anderson, Richard D.
APPLICANT: Ibraghimov, Oxana B.
TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: ME
COUNTRY: USA
ZIP: 03911
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,161A
FILING DATE: 16-SEP-93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/946,234
FILING DATE: 14-SEP-92
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M.
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: UIRF89-11A4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (207) 363-0558
TELEFAX: (207) 363-0528
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 895 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-123-161A-8

Query Match	37.1%;	Score 41.5;	DB 1;	Length 895;
Best Local Similarity	38.5%;	Pred. No. 1.4e+02;		
Matches 10; Conservative	2;	Mismatches 5;	Indels 9;	Gaps 1;

```

Qy      1 NSN-----RAHQGKHETGYGHT 17
          11      : 1111 : 11
Db      555 NSNSQLMYGLPDDSHVGKHEYFMHAT 580

```

RESULT 8
US-08-483-278-8
; Sequence 8, Application US/08483278
; Patent No. 5686073

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

?      COMPUTER READABLE FORM:
?      MEDIUM TYPE: Floppy disk
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: PatentIn Release #1.0, Version #1.25
?      CURRENT APPLICATION DATA: 00000000000000000000000000000000

```

Query Match	37.1%	Score 41.5	DB 1	Length 895
Best Local Similarity	38.5%	Pred. No. 1.4e+02		
Matches 10	Conservative 2	Mismatches 5	Indels 9	Gaps 1

QY 1 NSN-----RAHQCKHETGYCKKT 17
||| : |||| : |||
Db 555 NSNSQLMYGLPDDSHVCKHEYFMHAT 580

RESULT 9
US-09-134-001C-2962
; Sequence 2962, Application US/09134001C
; Patent No. 6380370

```

: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match	36.68;	Score 41;	DB 4;	Length 707;
Best Local Similarity	40.08;	Pred. No. 1.3e+02;		
Matches	6;	Conservative	3;	Mismatches 6;
				Indels 0;
				Gaps 0;

```
QY      1 NSNRAHQKHETCYH 15
          :|| | ||: |
Db      36 HSNHMHHDNHESHNH 50
```

RESULT 10
US-09-101-146-6
; Sequence 6, Application US/09101146
; Patent No. 6124125

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
3  COMPUTER: IBM PC
4  OPERATING SYSTEM: WINDOWS 95
5  SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
6  CURRENT APPLICATION DATA:

```

```

; TYPE: Amino acid
; TOPOLOGY: Linear
US-09-101-146-6

```

QY	8	GKHETYGHK	16
Db	20	GKHETLGHK	28

RESULT 11
US-09-101-146-1
Sequence 1, Application US/09101146
Patent No. 6124125
GENERAL INFORMATION:
APPLICANT: Dartmouth College, St. Vincents Institute of
APPLICANT: Medical Research, Kemp et al.
TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Milton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/101,146
FILING DATE: October 7, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PN7450
FILING DATE: 8 JAN 1996
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: DC-0050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515
TELEFAX: (856) 810-1454
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 345
TYPE: Amino acid
TOPOLOGY: linear
US-09-101-146-1

Query Match 35.7%; Score 40; DB 3; Length 345;
Best Local Similarity 77.8%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 GKHEHYGKH 16
DB 32 GKHEHYGKH 40

RESULT 12
US-09-320-878-8
Sequence 8, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BERTLACH, Melanie C.
APPLICANT: BERTLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
CURRENT FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/719,139
EARLIER FILING DATE: 1999-02-08

EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ. ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ. ID NO 8
LENGTH: 426
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-320-878-8

Query Match 35.7%; Score 40; DB 3; Length 426;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 AHOGKHEHYGKTP 18
DB 390 AHRLRETFGDPPT 403

RESULT 13
US-09-105-537-20
Sequence 20, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and plikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ. ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 20
LENGTH: 426
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-20

Query Match 35.7%; Score 40; DB 4; Length 426;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 AHOGKHEHYGKTP 18
DB 390 AHRLRETFGDPPT 403

RESULT 14
US-09-625-188-20
Sequence 20, Application US/09625188
Patent No. 6307037
GENERAL INFORMATION:
APPLICANT: No. 6307037artla AG
TITLE OF INVENTION: Fungal Target Genes and Methods
FILE REFERENCE: PB/5-31285P1
CURRENT APPLICATION NUMBER: US/09/625,188
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ. ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ. ID NO 20
LENGTH: 729
TYPE: PRT
ORGANISM: Ashbya gossypii
US-09-625-188-20

Query Match 35.7%; Score 40; DB 4; Length 729;
Best Local Similarity 44.4%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Oy 1 NSNRAHOGKHETYGHTP 18
 Db 18 NLRAFSSTENYGHMTP 35

RESULT 15

US-08-857-076-40
 ; Sequence 40, Application US/08857076C
 ; Patent No. 6225120
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruvkun, Gary
 ; APPLICANT: Kimura, Koutarou
 ; APPLICANT: Patterson, Garth
 ; APPLICANT: Ogg, Scott
 ; APPLICANT: Paradis, Suzanne
 ; APPLICANT: Tissenbaum, Heidi
 ; APPLICANT: Morris, Jason
 ; APPLICANT: Kowek, Allison
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
 ; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 00786/351001
 ; CURRENT APPLICATION NUMBER: US/08/857,076C
 ; CURRENT FILING DATE: 1997-05-15
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 40
 ; LENGTH: 796
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-08-857-076-40

Query Match 35.7%; Score 40; DB 4; Length 796;
 Best Local Similarity 28.6%; Pred. No. 2.1e+02;
 Matches 10; Conservative 1; Mismatches 8; Indels 16; Gaps 1;
 Oy 1 NSNRAHOGKH-----ETYGHTPY 19
 Db 375 NNNPQNHHNDISHPNHYSTDCGPNLYGFPPTY 409

Search completed: January 14, 2003, 18:17:06
 Job time : 5.22222 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:40:55 : Search time 12.1389 Seconds

(without alignments) 208.566 Million cell updates/sec

Title: US-09-834-794A-5

Sequence: 1 NSNRAGQKHETGKTPY 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	19	20	AAV29096 Human mammary asso
2	112	100.0	104	20	AAV29093 Human mammary asso
3	112	100.0	106	23	AAH47869 Mature placenta-de
4	112	100.0	127	19	AAH60649 Human DVLG-1 C-C c
5	112	100.0	127	20	AAV29092 Human mammary asso
6	112	100.0	127	21	AAH01450 Primate VIC Protei
7	112	100.0	127	22	AAH69175 Human G-protein co
8	112	100.0	127	23	AAH07728 Human mucosae-asso
9	112	100.0	127	23	AAH08268 Human mammary glan
10	112	100.0	127	23	AAH47868 Full length placen

11	112	100.0	147	22	AAH25479 Human protein sequ
12	107	95.5	127	21	AAH03001 Human growth facto
13	48	42.9	567	21	AAH18197 Plasmodium falci
14	48	42.9	692	22	AAH32897 Novel human secre
15	46	41.1	54	22	AAH51133 Propionibacterium
16	46	41.1	757	19	AAH71297 Lucilia cuprina ec
17	45	40.2	81	22	AAH62163 Propionibacterium
18	45	40.2	130	23	AAH23052 Human phospholip
19	45	40.2	192	21	AAH62924 HIV-1 non-subtype
20	45	40.2	427	23	AAH92637 Hericidially activ
21	45	40.2	500	13	AAH27743 Sequence transcrib
22	45	40.2	1035	22	AAH61120 Drosophila melanog
23	44	39.3	509	22	AAH57849 Drosophila melanog
24	44	39.3	509	22	AAH67227 Drosophila melanog
25	44	39.3	509	22	AAH67228 Drosophila melanog
26	44	39.3	630	14	AAH34664 Human 5HT transpor
27	44	39.3	630	16	AAH76074 Human 5HT transpor
28	44	39.3	630	22	AAH70190 Human 5HT transpor
29	43.5	38.8	115	23	AAH03049 Human ORFX protein
30	43.5	38.8	402	13	AAH24392 Sequence of the H1
31	43.5	38.8	654	21	AAH18162 Plasmodium falci
32	43	38.4	258	22	AAH02381 Novel human diagn
33	43	38.4	757	20	AAH01939 Ecdysone receptor
34	43	38.4	757	20	AAH28603 Ecdysone receptor
35	43	38.4	757	20	AAH67096 Ecr polypeptide su
36	43	38.4	774	22	AAH13676 Novel human diagn
37	42.5	37.9	164	22	AAH43501 Human polypeptide
38	42.5	37.9	325	22	AAH36327 Peptide #3833 enco
39	42.5	37.9	325	22	AAH36327 Peptide #3684 enco
40	42.5	37.9	325	22	AAH57092 Human brain expres
41	42.5	37.9	325	22	AAH69483 Human bone marrow
42	42.5	37.9	325	22	AAH17326 Peptide #3760 enco
43	42.5	37.9	325	22	AAH29825 Peptide #3862 enco
44	42.5	37.9	325	22	AAH05007 Peptide #3689 enco
45	42.5	37.9	325	23	AAH39116 Human peptide enco

ALIGNMENTS

RESULT 1
AAV29096 standard; peptide; 19 AA.
ID AAV29096:
XX
AC AAV29096:
XX
DT 29-SEP-1999 (first entry)
XX
DE Human mammary associated chemokine (MACK) protein fragment MACK C.
XX
KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
KW Inflammation; infection; mastitis; benign cystitis; hyperplasia;
KW mammary associated chemokine; MACK; epitope.
XX
OS Homo sapiens.
XX
PN WO936540-A1.
XX
PD 22-JUL-1999.
XX
PF 12-JAN-1999; 99WO-US00651.
XX
PR 09-JUL-1998; 98US-0092155.
PR 20-JAN-1998; 98US-0071899.
XX
PA (CODON) CODON DIAGNOSTICS LLC.
XX
PI Dyster LM, Frustaci JM, Papsidero LD;
XX WPI: 1999-458469/38.
XX A mammary associated chemokine and related polynucleotides, useful
PT for detection and treatment of breast disease, especially cancer

XX A mammary associated chemokine and related polynucleotides, useful
PT for detection and treatment of breast disease, especially cancer
PM

XX	
PS	Disclosure; Flg 1; 97pp; English

XX This sequence shows a chemokine protein, which is a placenta-derived
CC C-C chemokine (PLACC). The PLACC polypeptide or the chemokine-encoding
CC polynucleotide, are useful for modulating an inflammatory/immune
CC response, modulating the activity of PLACC and screening for modulators,
CC and in stimulating cell proliferation. Particularly these are useful
CC for treating asthma, adult respiratory distress syndrome, rheumatoid
CC arthritis, lupus erythematosus, psoriasis, osteoarthritis,
CC glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
CC Addison's Disease, Graves Disease, Crohn's Disease, irritable bowel
CC syndrome, atrophic gastritis, graft versus host disease, mastocytosis,
CC atherosclerosis or pancreatitis. Increasing the PLACC activity in the
CC subject is useful where the subject may be exhibiting signs of a
CC malignancy or infection or have an immunodeficiency.

XX
XX Sequence 106 AA:

SO Query Match 100.0%; Score 112; DB 23; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRAHOGKHETGKTPY 19
Db 88 NSNRAHOGKHETGKTPY 106

RESULT 4
AAW60649
ID AAW60649 standard; Protein: 127 AA.
XX
XX AAW60649;
XX
XX 12-OCT-1998 (first entry)
XX
XX Human DVIC-1 C-C chemokine.
XX
XX Human DVIC-1 C-C chemokine; cytokine; human; immune system;
XX
XX DVIC-1; DNAX Vlc-1; C-C chemokine; cytokine; human; immune system;
XX
XX cancer; cell proliferation; therapy; diagnosis.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..22
XX Protein 23..127
XX /label= Mat_protein
XX /note= "Claim 1"

XX
XX W09823750-A2.
XX
XX
XX 04-JUN-1998.
XX
XX
XX 26-NOV-1997; 97WO-US21092.
XX
XX
XX 05-DEC-1996; 96US-0761071.
XX
XX 27-NOV-1996; 96US-0031805.
XX
XX (SCHE) SCHERING CORP.
XX
XX Hedrick JA, Morales J, Vicari A, Zlotnik A;
XX
XX WPI; 1998-322730/28.
XX
XX N-PSDB; AAV38291.
XX
XX DVIC-1 and DGMCC chemokines - useful for developing products for
XX treating abnormal physiology or development, e.g. cancerous or
XX degenerative conditions
XX
XX Claim 1; Page 59-60; 71pp: English.
XX
XX This polypeptide comprises human DNAX Vlc-1 (DVIC-1), a novel C-C
XX chemokine, the mature portion of which is claimed. The amino acid

CC sequence was deduced from a cDNA clone (see AAV38291). An alternative
CC longer transcript (see AAW60652) for human DVIC-1 is also disclosed.
CC Also claimed is novel human DNAX Groin Wound expressed CC chemokine
CC (DGMCC) (see AAW60649) mature protein, as well as expression vectors
CC and host cells. DVIC-1 and DGMCC play a role in the regulation or
CC development of neuronal or haematopoietic cells, e.g. lymphoid
CC cells, which affect immunological responses. They can be used in
CC the treatment of conditions associated with abnormal physiology or
CC development, including abnormal proliferation, e.g. cancerous
CC conditions or degenerative conditions. Abnormal proliferation,
CC regeneration, degeneration, and atrophy may be modulated by
CC appropriate therapeutic treatment using products of the invention.
CC The products can also be used for detection, diagnosis and drug
CC screening.

XX
XX Sequence 127 AA:

SO Query Match 100.0%; Score 112; DB 19; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRAHOGKHETGKTPY 19
Db 109 NSNRAHOGKHETGKTPY 127

RESULT 5
AAV29092
ID AAV29092 standard; Protein: 127 AA.
XX
XX AAV29092;
XX
XX 29-SEP-1999 (first entry)
XX
XX Human mammary associated chemokine (MACK) protein.
XX
XX Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
XX inflammation; infection; mastitis; benign cystitis; hyperplasia;
XX mammary associated chemokine; MACK.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Peptide 1..23
XX Protein 24..127
XX /note= "signal peptide"
XX /note= "mature protein (AAV29093)"
XX Region 32..49
XX /note= "antigenic peptide epitope MACK A (AAV29094)"
XX MISC-difference 70
XX /label= unknown
XX /note= "encoded by NGA"
XX MISC-difference 91
XX /label= unknown
XX /note= "encoded by AAN"
XX Region 92..107
XX /note= "antigenic peptide epitope MACK B (AAV29095)"
XX Region 109..127
XX /note= "antigenic peptide epitope MACK C (AAV29096)"

XX
XX W09936540-A1.
XX
XX
XX 22-JUL-1999.
XX
XX
XX 12-JAN-1999; 99WO-US00651.
XX
XX 09-JUL-1998; 98US-0092155.
XX
XX 20-JAN-1998; 98US-0071899.
XX
XX (CODO-) CODON DIAGNOSTICS LLC.
XX
XX Dyster LM, Frustaci JM, Papsidero LD;

DR WPI: 1999-458469/38.
 DR N-PSDB: AAX89389, AAX89390.
 XX
 PT A mammary associated chemokine and related polynucleotides, useful
 PT for detection and treatment of breast disease, especially cancer
 XX
 PS Claim 4; Page 48; 76pp; English.
 XX
 CC The invention provides an isolated human chemokine, which is
 CC preferentially expressed in breast tissue or detected in breast milk. An
 CC antibody that recognizes the novel chemokine, or a chemokine-derived
 CC antigenic peptide, can be used to treat breast disease in a patient. A
 CC peptide, which binds to a cellular receptor for the chemokine, can also
 CC be used to treat breast disease. Antigenic peptides of the chemokine can
 CC be used to vaccinate patients against breast disease. The chemokine
 CC polynucleotide sequences and the chemokine protein can be detected in
 CC samples with primers, probes and antibodies using standard techniques.
 CC This is useful for detecting breast disease. Other breast diseases that
 CC may be treated or detected with the chemokine and its encoding
 CC polynucleotides include inflammations, infections, mastitis, benign
 CC cystitis, and benign hyperplasias as well as other malignancies. The
 CC present sequence represents the amino acid sequence of the human mammary
 CC associated chemokine (MACK) protein.
 CC
 SQ Sequence 127 AA;
 Query Match 100.0%; Score 112; DB 20; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NSNRAHQKHETGYGHKTPY 19
 DB 109 NSNRAHQKHETGYGHKTPY 127
 AAB01450
 ID AAB01450 standard; Protein: 127 AA.
 AC AAB01450;
 XX
 DT 20-OCT-2000 (first entry)
 DE Primate vic protein sequence.
 DE
 XX Cutaneous T-cell attracting chemokine; CTACK; skin; cell movement;
 KW migration; vasoactive intestinal contractor; VIC; GPR2; agonist;
 KW antagonist; antibody; immunological condition; mutein.
 XX
 OS Homo sapiens.
 XX
 PN WO200038713-A1.
 PD 06-JUL-2000.
 XX
 XX 23-DEC-1999; 99WO-US30819.
 PR 24-DEC-1998; 98US-0113858.
 PR 27-MAY-1999; 99US-0322580.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Wang W, Oldham ER, Soto H, Lul Y, Hudak SA, Homey B, Morales JM;
 PI Kellermann S, McEvoy LM, Zlotnik A;
 XX
 DR WPI: 2000-465633/40.
 DR N-PSDB: AAA47545.
 XX
 PT Modulating cell movement within the skin, useful for treating
 PT immunological skin conditions or diseases comprising administering T
 PT cell-attracting chemokine or vasoactive intestinal contractor chemokine
 PT agonists or antagonists
 XX

PS Example 3; Page 69; 79pp; English.
 XX
 CC Modulating movement of a cell within or to the skin of a mammal can
 CC be achieved by administering an antagonist or agonist of cutaneous T
 CC cell-attracting chemokine (CTACK) or vasoactive intestinal contractor
 CC (VIC) chemokine. The antagonist is selected from a mutein of natural
 CC CTACK or VIC, an antibody which neutralises CTACK or VIC or an
 CC antibody which block GPR2 ligand binding. The CTACK or VIC agonists
 CC or antagonists are useful for treating medical conditions or diseases
 CC associated with immunological conditions of the skin.
 CC
 SQ Sequence 127 AA;
 Query Match 100.0%; Score 112; DB 21; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NSNRAHQKHETGYGHKTPY 19
 DB 109 NSNRAHQKHETGYGHKTPY 127
 AAB69175
 ID AAB69175 standard; Protein: 127 AA.
 AC AAB69175;
 XX
 DT 26-APR-2001 (first entry)
 DE Human G-protein coupled receptor GPR27 protein SEQ ID NO:2.
 DE
 XX Human; G-protein coupled receptor; GPR27; antimicrobial; analgesic;
 KW cyostatic; antidiabetic; anorectic; antiaesthetic; antiparkinsonian;
 KW cardiant; hypertensive; hypotensive; diuretic; osteopathic; antilucer;
 KW cerebroprotective; antiallergic; antiemetic; tranquilizer; nootropic;
 KW antidepressant; neuroleptic; anticonvulsant; vaccine; gene therapy;
 KW infection; pain; cancer; diabetes; obesity; anorexia; bulimia; asthma;
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; vomiting;
 KW psychotic disorder; neurological disorder; dyslexia.
 XX
 OS Homo sapiens.
 XX
 PN WO200107482-A1.
 PD 01-FEB-2001.
 XX
 PF 21-JUL-2000; 2000WO-US19855.
 XX
 XX 27-JUL-1999; 99US-0361564.
 PR
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Testa TT, Vawter L;
 XX
 DR WPI: 2001-191359/19.
 DR N-PSDB: AAF59231.
 XX
 PT New G-protein coupled receptor GPR27 polypeptides and polynucleotides,
 PT useful as vaccines or for treating diseases, e.g. infections, pain,
 PT cancers, urinary retention, osteoporosis, stroke, psychotic and
 PT neurological disorders
 XX
 PS Claim 1; Page 29-30; 31pp; English.
 XX
 CC The present sequence represents a human G-protein coupled receptor,
 CC designated GPR27. GPR27 has antimicrobial, analgesic, cyostatic,
 CC antidiabetic, anorectic, antiaesthetic, antiparkinsonian, cardiant,
 CC hypertensive, hypotensive, diuretic, osteopathic, cerebroprotective,
 CC antilucer, antiallergic, antiemetic, tranquilizer, antidepressant,

CC neuroleptic, nootropic and anticonvulsant activities, and can be used
CC in producing vaccines and in gene therapy. The GPR27 protein and
CC polynucleotide are useful for treating certain diseases or as vaccines
CC against these diseases. These diseases include bacterial, fungal,
CC protozoan or viral (e.g. infections caused by HIV-1 or HIV-2) infections,
CC pain, cancers, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, myocardial infarction, stroke,
CC ulcers, allergies, benign prostatic hypertrophy, migraine, vomiting,
CC psychotic and neurological disorders (e.g. anxiety, schizophrenia, manic
CC depression, depression, delirium, dementia or mental retardation), or
CC dyskinesias (e.g. Huntington's disease or Gilles de la Tourette's
CC syndrome). The polynucleotide may also be used for chromosome
CC localisation studies, as tools for tissue expression studies, or as a
CC diagnostic reagent for detecting mutations in the associated gene.
CC
XX
SQ Sequence 127 AA;
Query Match 100.0%; Score 112; DB 22; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NSNRAHOGKHETYGKTPY 19
|||||
Db 109 NSNRAHOGKHETYGKTPY 127
RESULT 8
ABB07728
ID ABB07728 standard; Protein: 127 AA.
AC ABB07728;
DT 10-JUN-2002 (first entry)
DE Human mucosae-associated epithelial chemokine (MHC).
XX
XX Mucosae-associated epithelial chemokine; MHC; C-C chemokine receptor;
KW CCR3; CCR10; anti-inflammatory; cytosolic; immunomodulator; anti-viral;
KW antibacterial; chemokine; human.
XX
OS Homo sapiens.
XX
XX MO200214532-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001MO-US25734.
XX
XX 15-AUG-2000; 2000US-0638914.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Butcher EC, Kunkel EJ, Pan J, Soler-Ferran D;
XX
XX WPI: 2002-269204/31.
DR N-PSDB; ABL40453.
XX
XX Identifying modulators of mucosae-associated epithelial chemokine (MHC)
PT receptors 3 or 10 (CCR3/10), useful for treating inflammatory diseases,
PT comprises detecting formation of MHC-CCR3/10 complex or modulation of a
PT MHC-induced response -
XX
XX Example 1; Fig 1A; 92pp; English.
XX
XX The invention relates to identifying agents that inhibit or promote the
CC binding of a mammalian mucosae-associated epithelial chemokine (MHC) to
CC a mammalian C-C chemokine receptor 3 (CCR3) or 10 (CCR10). The method
CC involves: (a) detecting or measuring the formation of a complex between
CC the MHC, and the CCR3 or CCR10; or (b) determining the ability of the
CC test agent to inhibit or augment a MHC-induced response. An augmentation
CC of complex formation, relative to a control, is indicative that the agent

CC is a promoter. The method is useful for identifying modulators (e.g.
CC inhibitors or promoter) of MHC-induced functions of CCR3 and/or CCR10.
CC The inhibitors are useful for treating inflammatory diseases or
CC conditions in a subject, e.g. oral inflammatory condition (e.g. Sjogren's
CC syndrome or Behcet's syndrome), mastitis, chronic obstructive lung
CC disease, asthma, inflammatory bowel disease (e.g. Crohn's disease,
CC ulcerative colitis or celiac disease), Iga nephropathy or dermatitis
CC herpeticiformis. The promoters are useful for treating cancers (e.g. solid
CC tumours or cutaneous T cell lymphoma), neoplastic disease, retinopathy,
CC macular degeneration, bacterial infections, tuberculous leprosy, viral
CC infections, AIDS, neutropenias or bronchiectasis. The present sequence
CC represents the human MHC protein.
XX
XX
SQ Sequence 127 AA;
Query Match 100.0%; Score 112; DB 23; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NSNRAHOGKHETYGKTPY 19
|||||
Db 109 NSNRAHOGKHETYGKTPY 127
RESULT 9
ABB08268
ID ABB08268 standard; Protein: 127 AA.
AC ABB08268;
DT 20-MAY-2002 (first entry)
DE Human mammary gland enriched chemokine.
XX
XX Human; MHC; mammary gland enriched chemokine; chemokine; tumour; cancer;
KW cytosolic; anti-inflammatory; inflammation.
XX
OS Homo sapiens.
XX
XX Key location/Qualifiers
FT Peptide 1..19
FT Protein /label- Leader_sequence
FT Protein 20..127
XX
XX /label- Mature_MHC
XX
XX US2002009735-A1.
XX
XX 24-JAN-2002.
XX
XX 21-MAR-2001; 2001US-0813492.
XX
XX 23-MAR-2000; 2000US-191654P.
XX
XX (LABO/) LABOW M A.
PA (MICK/) MICKANIN C S.
PA (BHAT/) BHATIA U.
XX
PI Labow MA, Mickanin CS, Bhatia U;
XX
XX WPI: 2002-187776/24.
DR N-PSDB; ABA99025.
XX
XX Regulating tumour or adverse bodily reaction, involves providing
PT therapeutic composition comprising a mammary gland chemokine, and
PT providing the composition to the tumour or to the area of adverse
PT reaction -
XX
XX Claim 3; Fig 1; 11pp; English.
XX
XX The sequence represents human mammary gland enriched chemokine (MHC). The
CC invention relates to a novel method for regulating a tumour or
CC adverse bodily reaction, comprising providing a therapeutic composition
CC having a mammary gland chemokine polypeptide. The polypeptide of the

CC invention has cytostatic and antiinflammatory activity. The method of the
 CC invention is useful for regulating a tumour or adverse bodily reaction.
 CC The invention also provides a method useful for detecting a tumour using
 CC a probe comprising the polynucleotide or an antibody to the MEC. The
 CC adverse bodily reactions include cancer and inflammation.

XX Sequence 127 AA:

Query Match 100.0%; Score 112; DB 23; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRAHQKHETYGHKTPY 19
 |||||
 DB 109 NSNRAHQKHETYGHKTPY 127

RESULT 10

AAB47868 standard; Protein; 127 AA.

AC AAB47868;

DF 15-APR-2002 (first entry)

DE Full length placenta-derived C-C chemokine.

KW Chemokine; placenta-derived; C-C chemokine; PLACC; inflammation;
 KW Immune response; cell proliferation; asthma; malignancy; infection;
 KW adult respiratory distress syndrome; rheumatoid arthritis;
 KW lupus erythematosus; psoriasis; osteoarthritis; glomerulonephritis;
 KW osteoporosis; dermatomyositis; polymyositis; Addison's Disease;
 KW Graves Disease; Crohn's Disease; Irritable bowel syndrome;
 KW atrophic gastritis; graft versus host disease; myasthenia gravis;
 KW multiple sclerosis; autoimmune thyroiditis; ulcerative colitis;
 KW atherosclerosis; pancreatitis.

XX Homo sapiens.

OS Key Location/Qualifiers

FF Peptide 1..21

FT /label= Signal_peptide

FT Protein 22..127

FT /label= C-C_chemokine

PN WO200192301-A2.

PD 06-DEC-2001.

PF 23-MAY-2001; 2001WO-US16599.

PR 26-MAY-2000; 2000US-207578P.

PA (INDV) UNIV INDIANA ADVANCED RES & TECHNOLOGY.

PI Hromas RA;

PR WPI; 2002-154522/20.

DR N-PSDB; AAI72313.

XX New placenta-derived human C-C chemokine, useful for modulating an
 XX inflammatory or immune response, particularly for treating asthma,
 XX psoriasis, osteoarthritis, Graves Disease, Crohn's Disease or graft
 XX versus host disease

PS Claim 1; Page 90-91; 97pp; English.

CC This sequence shows a chemokine protein, which is a placenta-derived
 CC C-C chemokine (PLACC). The PLACC polypeptide or the chemokine-encoding
 CC polynucleotide, are useful for modulating an inflammatory/immune
 CC response, modulating the activity of PLACC and screening for modulators,
 CC and in stimulating cell proliferation. Particularly these are useful
 CC for treating asthma, adult respiratory distress syndrome, rheumatoid

CC arthritis, lupus erythematosus, psoriasis, osteoarthritis,
 CC glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
 CC Addison's Disease, Graves Disease, Crohn's Disease, Irritable bowel
 CC syndrome, atrophic gastritis, graft versus host disease, myasthenia
 CC gravis, multiple sclerosis, autoimmune thyroiditis, ulcerative colitis,
 CC atherosclerosis or pancreatitis. Increasing the PLACC activity in the
 CC subject is useful where the subject may be exhibiting signs of a
 CC malignancy or infection or have an immunodeficiency.

XX Sequence 127 AA:

Query Match 100.0%; Score 112; DB 23; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.7e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRAHQKHETYGHKTPY 19
 |||||
 DB 109 NSNRAHQKHETYGHKTPY 127

RESULT 11

AAM25479 standard; Protein; 147 AA.

AC AAM25479;

DF 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:994.

KW Human; cancer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; viricide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; anaemic;
 KW antiagregant; haemostatic; vulnerrary; antitumor; osteopathic; eczema;
 KW dermatological; antiallergic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antiangiogenic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.

XX Homo sapiens.

OS WO200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US35017.

PR 23-DEC-1999; 99US-0471275.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-457603/49.

DR N-PSDB; AAH99420.

XX Isolated human polynucleotides encoding polypeptides, useful for the
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection

PS Claim 20; Page 207; 1217pp; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antinflammatory; antirheumatic;

CC antitubercular; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; virocid; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antineoplastic; antidiabetic; haemostatic; vlnary;
 CC antitumor; osteoporosis; dermatological; antiallergic; antistatic;
 CC antidiabetic; cytotoxic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 encoding them can be used in gene therapy, antisense therapy and vaccine
 production. The proteins and polynucleotides are useful for screening for
 agonists or antagonists of a protein and for the treatment and diagnosis
 of disorders associated with the activity of a protein e.g. inflammation,
 rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

XX Sequence 147 AA;

Query Match 100.0%; Score 112; DB 22; Length 147;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGKTPY 19
 |||||
 Db 129 NSNRHOGKHETGKTPY 147

RESULT 12
 AAB03001 standard; Protein; 127 AA.

XX AAB03001;

DT 25-SEP-2000 (first entry)

XX Human growth factor related molecule GFRP-2.

XX Human GFRP-2; growth factor related molecule; breast tissue;
 KM breast tumour; CC chemokine; hTREC homologue; developmental disorder;
 KM cell proliferative disorder; immune disorder; reproductive disorder;
 KM cardiovascular disorder; bacterial infection; viral; fungal; parasitic;
 KM cancer; allergy; asthma; arteriosclerosis; therapy; diagnosis.

XX Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..19 /note- "Putative signal peptide"

FT Peptide 1..22 /note- "Putative signal peptide"

FT Region 30..74 /note- "This region has strong similarity to the CC chemokine consensus sequence"

FT Modified-site 39 /note- "Phosphorylated by protein kinase C"

FT Modified-site 78 /note- "N-glycosylated"

FT Modified-site 80 /note- "Phosphorylated by protein kinase C"

FT Modified-site 110 /note- "Phosphorylated by cAMP/cGMP-dependent protein kinase or protein kinase C"

XX WO200024774-A2.

XX 04-MAY-2000.

XX 28-OCT-1999; 99MO-US25458.

XX 28-OCT-1998; 98US-0181711.

PR 11-DEC-1998; 98US-0209547.
 PR 17-MAY-1999; 99US-0313457.
 XX
 PA (INCY-) INCYTE PHARM INC.

PI Tang YT, Yue H, Hillman JL, Corley NC, Guegler KJ, Baughn MR,
 PI Au-Young J;

XX WPI; 2000-350695/30.
 DR N-PSDB; AAA52456.

PT Human growth factor related molecule protein useful for the diagnosis
 PT and treatment of disorders associated with its activity including
 PT developmental, cell proliferative, immune, reproductive and
 PT cardiovascular disorders and infections -

XX Claim 1; Fig 2; 80pp; English.

CC This sequence represents human growth factor related molecule GFRP-2.

CC cDNA encoding GFRP-1 was initially identified in a breast tissue

CC cDNA library, and the present sequence is encoded by a consensus

CC GFRP-2 is probably a CC chemokine and has chemical and structural

CC homology with hTREC (20% identity). GFRP-2 and hTREC are both basic

CC proteins, having isoelectric points of 10.1 and 10.2, respectively.

CC GFRP-2 was found by Northern analysis to be expressed in both tumorous

CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),

CC nucleotides encoding them (AAA52455-AAA52458), GFRP agonists and

CC antagonists may be used to treat a wide variety of diseases associated

CC with increased or decreased expression or activity of GFRP proteins.

CC Conditions which may be treated include developmental disorders, cell

CC proliferative disorders (e.g., cancers), immune disorders (e.g.,

CC allergies, asthma), reproductive disorders (e.g., menstrual cycle

CC disorders) cardiovascular disorders (e.g., arteriosclerosis) and

CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP

CC proteins and nucleotides can be used in the diagnosis of such disorders.

XX Sequence 127 AA;

Query Match 95.5%; Score 107; DB 21; Length 127;
 Best Local Similarity 94.7%; Pred. No. 1.1e-09;
 Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NSNRHOGKHETGKTPY 19
 :|||

Db 109 DSNRHOHGHETGKTPY 127

RESULT 13

AAB18197 standard; Protein; 567 AA.

XX AAB18197;

DT 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:54.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;

XX antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.

XX WO200025728-A2.

XX 11-MAY-2000.

XX 05-NOV-1999; 99MO-US26796.

XX 05-NOV-1998; 98US-0107131.

XX (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
vv

PS Example 1; SEQ ID No 12328; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent *Propionibacterium* acnes immunogenetic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC *P. acnes* is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of *P. acnes* in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for *P. acnes* proteins. These antibodies can be used to
CC downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.

50 Sequence 54 AA;

Query Match	41.1%;	Score 46;	DB 22;	Length 54;
Best Local Similarity	46.7%;	Pred. NO. 3.6;		
Matches	7;	Conservative	3;	Mismatches 5;
				Indels 0;
				Gaps 0;

Search completed: January 14, 2003, 18:13:18
Job time : 13.1389 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:42:25 ; Search time 2.3111 Seconds

(without alignments)
287.144 Million cell updates/sec

Title: US-09-834-794a-4

Perfect score: 97

Sequence: 1 KNGKGNVCHRRKHGK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

112892

Total number of hits satisfying chosen parameters:

0

Minimum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	127	1	SY28_HUMAN
2	44	45.4	133	1	WN7B_ALOUV
3	43	44.3	130	1	SY28_MOUSE
4	42	43.3	295	1	UNG_MOUSE
5	42	43.3	603	1	SYD_AOUAE
6	42	43.3	1640	1	CO3_ONCMY
7	41	42.3	809	1	SYO_YEAST
8	41	42.3	1078	1	GYRB_SYNY3
9	41	42.3	1374	1	RNC_HUMAN
10	40.5	41.8	942	1	M3KE_MOUSE
11	40	41.2	89	1	ALB1_PIRANU
12	40	41.2	583	1	STS_HUMAN
13	40	41.2	606	1	ZG66_XENLA
14	40	41.2	509	1	NPRV_VIBPR
15	40	41.2	757	1	LOL4_MOUSE
16	40	41.2	1744	1	TANA_XENLA
17	39	40.2	69	1	YHDL_HAEIN
18	39	40.2	155	1	YHDL_HAEIN
19	39	40.2	235	1	OVOL_HUMAN
20	39	40.2	413	1	YVAC_CAEEL
21	39	40.2	427	1	FKH2_XENLA
22	39	40.2	756	1	LOLA_HUMAN
23	39	40.2	810	1	GYRA_BORBU
24	39	40.2	893	1	DAG1_MOUSE
25	39	40.2	895	1	DAG1_BOVIN
26	39	40.2	895	1	DAG1_HUMAN
27	39	40.2	895	1	DAG1_RABIT
28	39	40.2	2483	1	MPRI_MOUSE
29	39	40.2	5147	1	FAT_DROME
30	38.5	39.7	833	1	GYRA_BACHD
31	38	39.2	88	1	RS15_HAEIN
32	38	39.2	106	1	PER2_RHORN
33	38	39.2	244	1	RS6_BRATL

34	38	39.2	260	1	RS4_YARLI	059950 yarrowia 11
35	38	39.2	266	1	RS4_DICDI	P51403 dictyostell
36	38	39.2	283	1	VP40_MABVM	P35260 marburg vir
37	38	39.2	303	1	VP40_MABVP	003040 marburg vir
38	38	39.2	304	1	UNG_HUMAN	P13051 homo saplen
39	38	39.2	511	1	PUR9_BACHD	09Kf53 b ifunctio
40	38	39.2	697	1	Y4YR_RHISN	P55726 rhizobium s
41	38	39.2	1121	1	YE06_YEAST	P32644 saccharomyc
42	38	39.2	2491	1	MPRI_HUMAN	P11717 homo saplen
43	37	38.1	57	1	HIS1_HUMAN	P15515 homo saplen
44	37	38.1	97	1	HYPL_TRIRE	P52754 trichoderma
45	37	38.1	98	1	Y279_HAEIN	P43977 haemophilus

ALIGNMENTS

RESULT 1
SY28_HUMAN
ID SY28_HUMAN STANDARD; PRT; 127 AA.
AC Q9NRJ3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine A28 precursor (CCU28) (Mucosae-associated epithelial chemokine) (MCC) (CC128 protein).
DE SCYA28.
GN SCYA28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=20357357; PubMed=10781587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Honey B., Catron D., Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J., Kershenovich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.;
RT "Identification of a novel chemokine (CCL28), which binds CCR10 (GPR2).";
RT J. Biol. Chem. 275:22313-22323(2000).
RL [2]
RN [2]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RX MEDLINE=20432268; PubMed=10975800;
RA Pan J., Kunkel E.J., Gossler U., Lazarus N., Broadwell K., Vieira M.A., Genovese M.C., Butcher E.C., Soler D.;
RT "A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial cells in mucosal tissues.";
RT J. Immunol. 165:2943-2949(2000).
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang W., He L., Yuan Z., Wan T., Cao X.;
RT "A novel CC chemokine homology with TECK.";
RT Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
RL [4]
RP FUNCTION: CHEMOTACTIC ACTIVITY FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS. BINDS TO CCR3 AND CCR10 AND INDUCES CALCIUM MOBILIZATION IN A DOSE-DEPENDENT MANNER.
RL [5]
RP TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED BY EPITHELIAL CELLS OF DIVERSE TISSUES INCLUDING NORMAL AND PATHOLOGICAL COLON, SALIVARY GLAND, MAMMARY GLAND, TRACHEA AND RECTUM. ALSO FOUND IN PROSTATE, SPLEEN, THYROID, PSORIASIS SKIN AND IN LOWER LEVELS IN PERIPHERAL BLOOD LEUKOCYTES, SMALL INTESTINE, PEYER'S PATCHES, STOMACH AND NORMAL SKIN.
RL [6]
RP SIMILARITY: BELONGS TO THE INTERCINE BETA FAMILY (SMALL CYTOKINE C-C) (CHEMOKINE CC).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@lsb-slb.ch.
 CC -----
 DR EMBL: AF220210; AAF87205.1; -
 DR EMBL: AF266504; AAG16691.1; -
 DR EMBL: AF110384; AAG43193.1; -
 DR MIM: 605240; -
 DR InterPro: IPR000827; CC_chemkine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
 KW Cytokine; Chemotaxis; Signal; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 127 SMALL INDUCIBLE CYTOKINE A28.
 FT DISULFID 30 58 BY SIMILARITY.
 FT DISULFID 31 73 BY SIMILARITY.
 FT CARBOHYD 78 78 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 127 AA; 14280 MW; 3E855163A2CB62 CRC64;
 Query Match 100.0%; Score 97; DB 1; Length 127;
 Best Local Similarity 100.0%; Pred. No. 2.5e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KNGGNVCHRRKHGK 16
 DB 92 KNGGNVCHRRKHGK 107
 RESULT 2
 WNTB.ALOUV STANDARD; PRT; 123 AA.
 ID WNTB.ALOUV
 AC P28106;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE WNT-7B protein (Fragment).
 GN WNT-7B.
 OS Alopia vulpinus (Thresher shark).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphi; Galeoidea; Lamniformes; Alopidae;
 OC Alopas.
 OX NCBI_TaxID=7852;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92279273; PubMed-1534411;
 RA Sidor A.;
 RT "Diversification of the Wnt gene family on the ancestral lineage of vertebrates."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:5098-5102(1992).
 CC -1- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
 CC -1- SUBCELLULAR LOCATION: Possibly secreted and associates with the extracellular matrix.
 CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@lsb-slb.ch).
 CC -----
 DR EMBL: M91257; AAA8543.1; -
 DR InterPro: IPR000970; Wnt_gtrfactor.
 DR Pfam: PF00110; wnt; 1.
 DR SMART: SM00097; WNT1.1.
 DR PROSITE: PS00246; WNT1; PARTIAL.
 KW Developmental protein; Glycoprotein.
 FT NON_TER 1 1
 FT NON_TER 123 123

SQ SEQUENCE 123 AA; 14153 MW; F978FD1CED2DCE CRC64;
 Query Match 45.4%; Score 44; DB 1; Length 123;
 Best Local Similarity 50.0%; Pred. No. 3.8;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 4 KGNVCHRRKHGK 15
 DB 85 QGRICNRTSHHG 96
 RESULT 3
 SY28.MOUSE STANDARD; PRT; 130 AA.
 ID SY28.MOUSE
 AC 09JIL2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Small inducible cytokine A28 precursor (CCL28).
 GN SCYA28.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
 RC TISSUE-Kidney;
 RX MEDLINE-20357357; PubMed-10781587;
 RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Homey B., Catron D., Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J., Kershenovich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.;
 RT "Identification of a novel chemokine (CCL28), which binds CCR10 (GPR2)." J. Biol. Chem. 275:22313-22323(2000).
 RL J. Biol. Chem. 275:22313-22323(2000).
 CC -1- FUNCTION: CHEMOTACTIC FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS (BY SIMILARITY). BINDS TO CCR10 AND INDUCES CALCIUM MOBILIZATION IN A DOSE-DEPENDENT MANNER.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN TESTIS, EPITHELIAL CELLS OF NORMAL COLON, KIDNEY, PEYER'S PATCHES, LYMPH NODES. ALSO FOUND IN LOWER LEVELS IN BRAIN, SPLEEN AND LUNG.
 CC -1- SIMILARITY: BELONGS TO THE INTERKINE BETA FAMILY (SMALL CYTOKINE C-C) (CHEMOKINE CC).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@lsb-slb.ch).
 CC -----
 DR EMBL: AF220238; AAF87206.1; -
 DR MGD: MGI:1861731; Scya28.
 DR InterPro: IPR000827; CC_chemkine_sml.
 DR InterPro: IPR001811; Chemokine_IL8.
 DR SMART: SM00199; SCY; 1.
 DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.
 KW Cytokine; Chemotaxis; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT CHAIN 17 130 SMALL INDUCIBLE CYTOKINE A28.
 FT DISULFID 30 58 BY SIMILARITY.
 FT DISULFID 31 73 BY SIMILARITY.
 FT CARBOHYD 78 78 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 130 AA; 14570 MW; 6F3B909A4F97013 CRC64;
 Query Match 44.3%; Score 43; DB 1; Length 130;
 Best Local Similarity 66.7%; Pred. No. 5.8;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 KNGGNVCHRRK 12
 DB 92 KNGRENVCGRK 103

RESULT 4
 ID UNG_MOUSE STANDARD: PRT: 295 AA.
 AC P97931: P97509: (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Uracil-DNA glycosylase, mitochondrial precursor (EC 3.2.2.-) (UDG).
 GN UNG OR UNG1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97311407; PubMed=9168124;
 RA Swendsen P.C., Yee H.A., Winkfein R.J., van de Sande J.H.;
 RT "The mouse uracil-DNA glycosylase gene: Isolation of cDNA and genomic
 clones and mapping ung to mouse chromosome 5.";
 RL Gene 189:175-181(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97169285; PubMed=9016624;
 RA Nilsson H., Solum K., Haug T., Krokan H.E.;
 RT "Nuclear and mitochondrial uracil-DNA glycosylases are generated by
 alternative splicing and transcription from different positions in
 the UNG gene.";
 RL Nucleic Acids Res. 25:750-755(1997).
 CC -1- FUNCTION: EXCISES URACIL RESIDUES FROM THE DNA WHICH CAN ARISE
 AS A RESULT OF MISINCORPORATION OF DUMP RESIDUES BY DNA
 POLYMERASE OR DUE TO DEMINATION OF CYTOSINE.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AND MITOCHONDRIAL (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE URACIL-DNA GLYCOSYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U55040; AAC39511.1; -;
 DR EMBL: U55041; AAC53197.1; -;
 DR EMBL: X59018; CA67489.1; -;
 DR HSSP: P13051; 1AKZ.
 DR MOD: MGI:109352; ung.
 DR InterPro: IPR005122; UDNAL_glycos.
 DR InterPro: IPR002043; U_DNA_glycosylase.
 DR InterPro: IPR003249; U_glycosyl.
 DR Pfam: PF03167; UDG; 1.
 DR Pfam: PF001589; U_glycosyl; 1.
 DR TIGRfam: TIGR00628; ung; 1.
 DR PROSITE: PS00130; U_DNA_GLYCOSYLASE; 1.
 DR DNA Repair: Hydrolyase; Glycosidase; Nuclear protein; Mitochondrion;
 KM Transil peptide.
 FT TRANSIT 1 58 MITOCHONDRION (POTENTIAL).
 FT CHAIN 59 295 URACIL-DNA GLYCOSYLASE.
 FT ACT_SITE 136 136 GENERAL BASE (BY SIMILARITY).
 FT CONFLICT 266 266 H -> Y (IN REF. 2).
 SQ SEQUENCE 295 AA; 33054 MW; 7E6E56DEC55B851 CRC64;

Query Match 43.3%; Score 42; DB 1; Length 295;
 Best Local Similarity 63.6%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 KGNVCHRRKHH 14
 DB 243 KGSYIDKRRHH 253

RESULT 5
 ID SYD_AQUAE STANDARD: PRT: 603 AA.
 AC 067589;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 DE (AspRS).
 GN ASPS OR AQ.1677.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 OC Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) -> AMP +
 diphosphate + L-aspartyl-tRNA(Asp).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000750; AAC07548.1; -;
 DR HSSP: P36419; 1EFM.
 DR InterPro: IPR002106; AATRNA_ligaset1.
 DR InterPro: IPR004524; ASPS_bact.
 DR InterPro: IPR004115; GAD_dom.
 DR InterPro: IPR004364; tRNA-synt_2.
 DR InterPro: IPR002312; tRNA-synt_2.
 DR InterPro: IPR004365; tRNA-anti1.
 DR Pfam: PF00152; tRNA-synt_2; 2.
 DR Pfam: PF01336; tRNA-anti1; 1.
 DR Pfam: PF02938; GAD; 1.
 DR PRINTS: PR01042; TRNASYNTHASP.
 DR TIGRfam: TIGR00459; asps_bact; 1.
 DR PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KM Complete proteome.
 SQ SEQUENCE 603 AA; 69729 MW; 9DFEBD840C8DC1C CRC64;

Query Match 43.3%; Score 42; DB 1; Length 603;
 Best Local Similarity 85.7%; Pred. No. 34;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 9 HRRKHHG 15
 DB 33 HRRKHHG 39

RESULT 6
 ID CO3_ONCMY STANDARD: PRT: 1640 AA.
 AC P98093;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE Complement C3-1 [Contains: C3a anaphylatoxin] (Fragment).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RX TISSUE=Liver;
 RC MEDLINE=94065166; PubMed=8245455;
 RA Lambris J.D., Lao Z., Pang J., Alsenz J.;
 RT Third component of trout complement. cDNA cloning and conservation
 of functional sites.";
 RL J. Immunol. 151:6123-6134(1993).
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
 COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
 REACTION IN BOTH CLASSICAL AND ALTERNATIVE PATHWAYS.
 CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
 CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
 CC RESIDUES, FORMING TWO CHAINS, BETA 6 ALPHA, LINKED BY A DISULFIDE
 CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN.
 CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA
 CC CHAIN).
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
 CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L24433; AAB05029.1; ALT_INIT.
 DR HSSP: P01024; IC3D.
 DR InterPro: IPR002890; A2M_N.
 DR InterPro: IPR000020; Anaphylatoxin.
 DR InterPro: IPR001599; Macroglobin2.
 DR InterPro: IPR001134; Netrin_C.
 DR Pfam: PF00207; A2M; 1.
 DR Pfam: PF01759; NTR; 1.
 DR Pfam: PF01821; ANATO; 1.
 DR Pfam: PF01835; A2M_N; 1.
 DR ProDom: PD003264; Anaphylatoxin; 1.
 DR SMART: SM00104; ANATO; 1.
 DR PROSITE: PS00477; ALPHA-2-MACROGLOBULIN; 1.
 DR PROSITE: PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE: PS01178; ANAPHYLATOXIN_2; 1.
 DR Complement pathway; Complement alternate pathway; Plasma;
 KW Inflammatory response; Glycoprotein.
 FT NON_TER 1 1
 FT CHAIN 1 1640
 FT CHAIN 1 642
 FT CHAIN 1640
 FT PEPTIDE 647 722
 FT CHAIN 722
 FT PEPTIDE 723 1640
 FT CHAIN 1640
 FT PEPTIDE 723 931
 FT PEPTIDE 932 1278
 FT PEPTIDE 1278
 FT PEPTIDE 1279 1295
 FT PEPTIDE 1295
 FT SITE 722 723
 FT SITE 931 932
 FT SITE 1278 1279
 FT SITE 1295 1296
 FT SITE 1296 703
 FT DOMAIN 668
 FT DOMAIN 1420 1430
 FT DISULFID 536 797
 FT DISULFID 603 638
 FT DISULFID 668 695
 FT DISULFID 669 702
 FT DISULFID 669 702

FT DISULFID 682 703
 BY SIMILARITY.
 FT DISULFID 853 1488
 BY SIMILARITY.
 FT DISULFID 1079 1135
 BY SIMILARITY.
 FT DISULFID 1335 1464
 BY SIMILARITY.
 FT DISULFID 1481 1486
 BY SIMILARITY.
 FT DISULFID 1493 1563
 BY SIMILARITY.
 FT DISULFID 1510 1638
 BY SIMILARITY.
 FT DISULFID 1614 1623
 BY SIMILARITY.
 FT CARBOHYD 164 164
 N-LINKED (GLCNAC. . .).
 FT THIOLEST 988 991
 BY SIMILARITY.
 SO SEQUENCE 1640 AA; 182104 MW; 0965B4FAF1E87812 CRC64;
 Query Match 43.3%; Score 42; DB 1; Length 1640;
 Best Local Similarity 53.8%; Pred. NO. 84;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 OY 1 KNGKGVNCHRRKH 13
 DB 284 KDGKGVACLKKEH 296
 RESULT 7
 SYO_YEAST STANDARD; PRT; 809 AA.
 ID SYO_YEAST
 AC P13188; Q12005;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glutamyl-tRNA synthetase (EC 6.1.1.18) (Glutamine--tRNA ligase)
 DE (GLNRS).
 GN GLN OR YOR168W OR C3601.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RX MEDLINE=87280149; PubMed=3301841;
 RA Luderer S.W., Schimmel P.;
 RT "Gene for yeast glutamine tRNA synthetase encodes a large
 RT amino-terminal extension and provides a strong confirmation of the
 RT signature sequence for a group of the aminoacyl-tRNA synthetases";
 RL J. Biol. Chem. 262:10801-10806(1987).
 RN [12]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97127828; PubMed=8972579;
 RX Madania A., Poch O., Tarassov I.A., Winsor B., Matlin R.P.;
 RT "Analysis of a 22,956 bp region on the right arm of Saccharomyces
 RT cerevisiae chromosome XV";
 RL Yeast 12:1563-1573(1996).
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamine + tRNA(Gln) = AMP +
 CC diphosphate + L-glutamyl-tRNA(Gln).
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M29184; AAA34646.1; -;
 DR EMBL: M29185; -; NOT_ANNOTATED_CDS.
 DR EMBL: U55021; AAB47415.1; -;
 DR EMBL: Z75076; CAA9374.1; -;
 DR PIR: A28494; SYBYOT.
 DR HSSP: P00962; IGRN.
 DR SGD: S0005694; GLN4.
 DR InterPro: IPR004514; GLN.
 DR InterPro: IPR000924; Glu_tRNA-synt_1c.
 DR InterPro: IPR001412; tRNA-synt_1.

DR Pfam: PF00749; tRNA-synt_1c; 1.
 DR TIGRfams: TIGR00440; gins1; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT SITE 258 268 "HIGH" REGION.
 FT SITE 495 498 "KMSKS" REGION.
 FT BINDING 498 498 ATP (B) SIMILARITY).
 FT CONFLICT 179 179 G -> Q (IN REF. 1).
 SQ SEQUENCE 809 AA; 93132 MW; C7AB13D02BC483F6 CRC64;
 Query Match 42.3%; Score 41; DB 1; Length 809;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 1 KNGKGVNCH 9
 Db 339 KNGKGVNCH 347
 RESULT 8
 ID GYRB_SYNY3 STANDARD; PRT; 1078 AA.
 AC P77966;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA gyrase subunit B (EC 5.99.1.3) [Contains: ssp gyrb intein].
 GN GYRB OR SL2005.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Keneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Natsu K., Okumura S.,
 RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90908; BAA17720.1; -;
 DR HSSP: P06983; 1A56;
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR002288; DNA_gyraseB_C.
 DR InterPro: IPR001241; DNA_topoisoi.
 DR InterPro: IPR002936; DNAPrim_toprim.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR InterPro: IPR002711; HHN.

DR InterPro: IPR003586; Hedgehog_hintc.
 DR InterPro: IPR003587; Hedgehog_hintN.
 DR InterPro: IPR002203; Intein.
 DR Pfam: PF00204; DNA_gyraseB_1.
 DR Pfam: PF00986; DNA_gyraseB_C; 1.
 DR Pfam: PF01751; Toprim; 1.
 DR Pfam: PF01844; HHN; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR PRINTS: PR00418; TP12FAMILY.
 DR ProDom: PD000616; DNA_topoisoi1; 1.
 DR ProDom: PD149633; DNA_gyraseB_C; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00305; HintC; 1.
 DR SMART: SM00306; HintC; 1.
 DR SMART: SM0433; TOP2c; 1.
 DR TIGRfams: TIGR01059; gyrb; 1.
 DR PROSITE: PSS0818; INTEIN_C_TER; 1.
 DR PROSITE: PSS0817; INTEIN_N_TER; 1.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KW Isomerase; Topoisomerase; ATP-binding; Autocatalytic cleavage;
 KM Protein splicing; Complete proteome.
 FT CHAIN 1 436
 FT CHAIN 1 436
 FT CHAIN 437 871
 FT CHAIN 872 1078
 SQ SEQUENCE 1078 AA; 122819 MW; 6CA02586DEFA607B CRC64;
 Query Match 42.3%; Score 41; DB 1; Length 1078;
 Best Local Similarity 46.7%; Pred. No. 82;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 Oy 1 KNGKGVNCHRRKHHG 15
 Db 462 KNGKGVNCHRRKHHG 476
 RESULT 9
 ID RNC_HUMAN STANDARD; PRT; 1374 AA.
 AC Q9NBR4; Q9NBR4; Q9Y2V9; Q9Y4Y0;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribonuclease III (EC 3.1.26.3) (RNase III) (P241).
 GN RNASE3L OR RN3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20538440; PubMed=10948199;
 RA Wu H., Xu H., Miraglia L.J., Crooke S.T.,
 RT "Human RNase III is a 160-kDa protein involved in preribosomal RNA
 RT processing.";
 RL J. Biol. Chem. 275:36957-36965(2000).
 RN [2]
 RP SEQUENCE OF 166-613 FROM N.A. (ISOFORM 2).
 RC TISSUE=Colon; PubMed=10976766;
 RX MEDLINE=20431278; PubMed=10976766;
 RA Gunther M., Lathier M., Brisson O.,
 RT "A set of proteins interacting with transcription factor Sp1
 RT identified in a two-hybrid screening.";
 RL Mol. Cell. Biochem. 210:131-142(2000).
 RN [3]
 RP SEQUENCE OF 603-1374 FROM N.A.
 RC TISSUE=Embryo;
 RA Isegawa T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Magatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,

OY 3 GK-GNCHRRKKHK 16
 DB 127 GKMAVCRGRGRK 141
 RESULT 11
 ALBI_PHAU STANDARD; PRT; 89 AA.
 ID ALBI_PHAU STANDARD; PRT; 89 AA.
 AC 09FRT8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Albumin 1 precursor (P1) [contains: PALA; Leginsulin (PALB)]
 DE (Fragment).
 GN LEG.
 OS Phascolus aureus (Mung bean) (Vigna radiata).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OX NCBI_TaxID=3916;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Blackmappe; TISSUE-Leaf;
 RA Sakita M., Takeoka M., Hirano H.;
 RT Leginsulin, a plant 4-kDa peptide with insulin-like function."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: LEGINSULIN BINDS TO BASIC 7S GLOBULIN (BG) AND
 CC STIMULATES ITS PHOSPHORYLATION ACTIVITY (BY SIMILARITY).
 CC -1- PTM: THREE DISULFIDE BONDS ARE PROBABLY PRESENT.
 CC -1- PTM: THE C-TERMINAL GLYCINE MAY BE REMOVED FROM LEGINSULIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB052881; BAB19938.1;
 KM Seed storage protein; Albumin; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 1
 FT CHAIN 2 38 POTENTIAL.
 FT PROPEP 39 46 LEGINSULIN (BY SIMILARITY).
 FT CHAIN 47 >89 POTENTIAL.
 FT NON_TER 89 89 PALA (POTENTIAL).
 SO SEQUENCE 89 AA; 9711 MW; 4ADEB97970831358 CRC64;
 Query Match 41.2%; Score 40; DB 1; Length 89;
 Best Local Similarity 50.0%; Pct. No. 12;
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 OY 1 KNGKGNVCHRRKKH 14
 DB 61 KKGSGNFCARYPNH 74
 RESULT 12
 STS_HUMAN STANDARD; PRT; 583 AA.
 ID STS_HUMAN STANDARD; PRT; 583 AA.
 AC P08842;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Steryl-sulfatase precursor (EC 3.1.6.2) (steroid sulfatase) (steryl-
 DE sulfatase sulfohydrolase) (Arylsulfatase C) (ASC).
 GN STS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=89340479; PubMed=2668275;
 RA Stein C., Hille A., Seidel J., Rijhout S., Wahed A., Schmidt B.,
 RA Geuze H., von Figura K.;
 RT "Cloning and expression of human steroid-sulfatase. Membrane
 RT topology, glycosylation, and subcellular distribution in BHK-21
 RT cells."
 RL J. Biol. Chem. 264:13865-13872(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87187642; PubMed=3032454;
 RA Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T.,
 RA Shapiro L.J.;
 RT "Cloning and expression of steroid sulfatase cDNA and the frequent
 RT occurrence of deletions in STS deficiency: implications for X-Y
 RT interchange."
 RL Cell 49:443-454(1987).
 RN [3]
 RP SEQUENCE OF 134-174 AND 461-583 FROM N.A.
 RX MEDLINE=89077541; PubMed=3203382;
 RA Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L.,
 RA Neiswanger K., Shapiro L.J.;
 RT "The human X-linked steroid sulfatase gene and a Y-encoded
 RT pseudogene: evidence for an inversion of the Y chromosome during
 RT primate evolution."
 RL Cell 55:1123-1135(1988).
 RN [4]
 RP SEQUENCE OF 22-45.
 RC TISSUE-Liver;
 RX MEDLINE=89352671; PubMed=2765556;
 RA Kawano J.-I., Kotani T., Ohtaki S., Minamino N., Matsuo H., Ohtsuka T.,
 RA Aikawa E.;
 RT "Characterization of rat and human steroid sulfatases."
 RL Biochim. Biophys. Acta 997:199-205(1989).
 RN [5]
 RP VARIANTS XLI LEU-361; ARG-372 AND TYR-446.
 RX MEDLINE=9210784; PubMed=1539590;
 RA Basler E., Grompe M., Parenti G., Yates J., Ballabio A.;
 RT "Identification of point mutations in the steroid sulfatase gene of
 RT three patients with X-linked ichthyosis."
 RL Am. J. Hum. Genet. 50:483-491(1992).
 CC -1- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO ESTROGENS
 CC DURING PREGNANCY.
 CC -1- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate +
 CC H(2)O = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANE. THE SEQUENCE SHOWS
 CC SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE
 CC PROTEIN IN THE MITOCHONDRIAL MEMBRANE.
 CC -1- DISEASE: DEFECTS IN STS ARE A CAUSE OF A VISIBLE PHENOTYPE OF
 CC SCALY SKIN, X-LINKED ICHTHYOSIS (XLI).
 CC -1- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04964; AAA60597.1;
 DR EMBL; M16505; AAA60596.1;
 DR EMBL; M23945; AAA60598.1;
 DR EMBL; M23556; AAA60599.1;
 DR PIR; A32641; A32641.
 DR PIR; A32641; A32641.
 DR PIR; A25961; A25961.
 DR PIR; S05415; S05415.
 DR HSSP; P15848; IFSU.
 DR Genew; HGNC:11425; STS.
 DR MIM; 308100;
 DR InterPro; IPR000917; Sulfatase.
 DR Pfam; PF00884; Sulfatase; 1.

DR PROSITE: PS00523; SULFATASE_1; 1.
 DR PROSITE: PS00149; SULFATASE_2; 1.
 KW Hydrolyase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
 KW Pregnancy; Signal; Disease mutation.
 FT SIGNAL 1 21
 FT CHAIN 22 583
 FT MOD_RES 75 75
 FT 2-AMINO-3-OXOPROPIONIC ACID (BY
 SIMILARITY).
 FT ACT_SITE 136 136
 FT TRANSMEM 185 206
 FT TRANSMEM 213 234
 FT CARBOHYD 47 47
 FT CARBOHYD 259 259
 FT CARBOHYD 333 333
 FT CARBOHYD 459 459
 FT VARIANT 341 341
 FT VARIANT 372 372
 FT VARIANT 446 446
 FT CONFLICT 23 23
 FT SEQUENCE 583 AA; 65492 MW; 74746AFA9D21A0A6 CRC64;

Query Match
 Best Local Similarity 41.2%; Score 40; DB 1; Length 583;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 NCKGNCVCHRRKHHG 15
 DB 142 HSKTDFCHHPLHNG 155

RESULT 13
 ZG66_XENLA STANDARD: PRT; 606 AA.
 ID 2G66_XENLA
 AC P18733;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Gastrula zinc finger protein XLCGP66.1 (Fragment).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE OF 1-295 FROM N.A.
 RX MEDLINE=89345612; PubMed=2503827;
 RA Knoechel W., Poetling A., Koester M., el Baradi T., Metfield W.,
 Boumestier T., Pieler T.;
 RT "Evolutionary conserved modules associated with zinc fingers in
 Xenopus laevis";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6097-6100(1989).
 RN [2]
 RP SEQUENCE OF 240-606 FROM N.A.
 RX MEDLINE=90040698; PubMed=2509712;
 RA Metfield W., el Baradi T., Mentzel H., Pieler T., Koester M.,
 Poetling A., Knoechel W.;
 RT "Second-order repeats in Xenopus laevis finger proteins";
 RL J. Mol. Biol. 208:639-659(1989).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M25873; AAA50020.1; -;
 DR PIR; H33282; H33282;
 DR PIR; S06582; S06582;

DR HSSP; P07248; 2ADR.
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR000822; znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 12.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00355; znf_C2H2; 11.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 11.
 DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 12.
 KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
 FT ZN_FING 273 295
 FT ZN_FING 300 322
 FT ZN_FING 328 350
 FT ZN_FING 384 407
 FT ZN_FING 413 435
 FT ZN_FING 441 464
 FT ZN_FING 470 492
 FT ZN_FING 498 521
 FT ZN_FING 527 549
 FT ZN_FING 555 578
 FT ZN_FING 584 606
 FT NON_TER 606
 FT SEQUENCE 606 AA; 69087 MW; 518755B9F1DC2FD CRC64;

Query Match
 Best Local Similarity 41.2%; Score 40; DB 1; Length 606;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 CHRRKHHGK 16
 DB 290 CHQKTHKCK 298

RESULT 14
 NERV_VIBPR STANDARD: PRT; 609 AA.
 ID NERV_VIBPR
 AC O00971;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Aneurismal protease precursor (EC 3.4.24.25) (Vibriolysin)
 DE (Aeromonolysin).
 GN NERV.
 OS Vibrio proteolyticus (Aeromonas proteolytica).
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
 OX NCBI_TaxID=671;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 197-215.
 RX MEDLINE=92201689; PubMed=1551587;
 RA David V.A., Deutch A.H., Sioma A., Pawlyk D., Alty A., Durham D.R.;
 RT "Cloning, sequencing and expression of the gene encoding the
 RT extracellular neutral protease, vibriolysin, of Vibrio
 RT proteolyticus";
 RL Gene 112:107-112(1992).
 CC -1- FUNCTION: EXTRACELLULAR ZINC METALLOPROTEASE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage of bonds with bulky
 CC hydrophobic groups in P2 and P1.
 CC -1- COFACTOR: BINDS 1 ZINC ION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M64809; AAA27548.1; -;
 DR PIR; JTO903; JTO903;
 DR HSSP; P14756; IEZM.
 DR MEROPS; M04.003; -;
 DR InterPro; IPR005075; Pep_M4_propep.

DR InterPro: IPR001570; Peptidase_M4.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01447; Peptidase_M4; 1.
DR Pfam: PF02868; Peptidase_M4_C; 1.
DR Pfam: PF03413; Pep_M4_Proped; 1.
DR PRINTS: PR00730; THERMOLYSIN.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Calcium; Zinc; Zymogen; Signal.
FT SIGNAL 1 24
FT PROPEP 25 197 POTENTIAL.
FT CHAIN 198 609 NEUTRAL PROTEASE.
FT METAL 343 343 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 344 344 BY SIMILARITY.
FT METAL 347 347 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 367 367 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 426 426 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 609 AA; 66362 MW; 8PAB1798D737FCA2 CRC64;
Query Match 41.2%; Score 40; DB 1; Length 609;
Best Local Similarity 50.0%; Pred. No. 69;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 NGKGNVCHRRKHHG 15
DB 64 NGKVKVRYQYTHHG 77
RESULT 15
LOLA_MOUSE
ID LOLA_MOUSE STANDARD; PRT: 757 AA.
AC Q924C6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lysyl oxidase homolog 4 precursor (EC 1.4.3.-) (Lysyl oxidase-like
protein 4) (Lysyl oxidase related protein C).
GN LOXL4 OR LOXC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21316447; PubMed=11292829;
RA Ito H., Akiyama H., Iguchi H., Iyama K.I. K., Miyamoto M., Ohsawa K.,
Nakamura T.;
RT "Molecular cloning and biological activity of a novel lysyl oxidase-
related gene expressed in cartilage.";
RL J. Biol. Chem. 276:24023-24029(2001).
CC -1- FUNCTION: May modulate the formation of a collagenous
extracellular matrix.
CC -1- COFACTOR: COPPER AND LYSINE TYROSYLQUINONE (LTKO) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).
CC -1- PTM: THE LYSINE-TYROSYLQUINONE IS GENERATED BY THE MODIFICATION OF
A TYROSYL RESIDUE CROSSLINKED TO THE EPSILON-AMINO GROUP OF A
LYSINE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LYSYL OXIDASE FAMILY.
CC -1- SIMILARITY: CONTAINS 4 SRCR DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF338440; AAK71933.1; -
DR MGD: MGI:1914823; Lox14.
DR InterPro: IPR001695; Lysyl_oxidase.
DR InterPro: IPR001190; Srcr_receptor.
DR Pfam: PF01186; Lysyl_oxidase; 1.
DR Pfam: PF00530; SRCR; 4.

DR ProDom: PD013887; Lysyl_oxidase; 1.
DR PROSITE: PS00926; LYSYL_OXIDASE; FALSE_NEG.
DR PROSITE: PS00420; SRCR_1; 1.
DR PROSITE: PS50287; SRCR_2; 4.
KW Oxidoreductase; Copper; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 25
FT CHAIN 26 757 POTENTIAL.
FT DOMAIN 33 134 LYSYL OXIDASE HOMOLOG 4.
FT DOMAIN 160 288 SRCR 1.
FT DOMAIN 312 412 SRCR 2.
FT DOMAIN 422 530 SRCR 3.
FT DOMAIN 534 737 SRCR 4.
FT METAL 612 612 LYSYL-OXIDASE LIKE.
FT METAL 614 612 COPPER (POTENTIAL).
FT METAL 616 616 COPPER (POTENTIAL).
FT MOD_RES 639 639 COPPER (POTENTIAL).
FT MOD_RES 639 639 CROSS-LINKED TO TYROSYLQUINONE (BY
SIMILARITY).
FT MOD_RES 675 675 TYROSYLQUINONE (BY SIMILARITY).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 757 AA; 84705 MW; D9861368F63B7B2A CRC64;
Query Match 41.2%; Score 40; DB 1; Length 757;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 5 GNVCHRRKHHG 15
DB 130 GNVCHPRRHG 140

Search completed: January 14, 2003, 18:13:51
Job time : 4.3111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:15:35 ; Search time 2.13333 Seconds

(without alignments)
145.508 Million cell updates/sec

Title: US-09-834-794a-4

Perfect score: 1 KNGKGNVCHRRKHGK 16

Sequence: 1 KNGKGNVCHRRKHGK 16

Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*

2: /cgn2_6/ptodata/2/pubpaa/PC7_NEW_PUB pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*

6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*

7: /cgn2_6/ptodata/2/pubpaa/PC7US_PUBCOMB pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*

11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*

12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*

13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*

14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	97	100.0	16	US-09-834-795a-4
2	97	100.0	104	US-09-834-795a-2
3	97	100.0	127	US-09-898-751a-6
4	97	100.0	127	US-09-813-492-2
5	97	100.0	127	US-09-834-795a-1
6	97	100.0	127	US-09-831-381a-2
7	43	44.3	130	US-09-898-751a-10
8	41.5	42.8	129	US-09-950-933a-43
9	41	42.3	1374	US-09-900-425a-2
10	40	41.2	166	US-09-738-626-6282
11	40	41.2	492	US-09-864-761-42897
12	40	41.2	757	US-09-823-038a-52
13	39.5	40.7	313	US-09-764-864-1435
14	39	40.2	83	US-09-764-847-504
15	39	40.2	147	US-09-798-042-101
16	39	40.2	170	US-10-067-422-14
17	39	40.2	185	US-09-159-469-10
18	39	40.2	185	US-09-798-042-10
19	39	40.2	573	US-10-067-422-10

20 39 40.2 756 10 US-09-870-110-2 Sequence 2, Appl1

21 39 40.2 756 10 US-09-924-946-2 Sequence 2, Appl1

22 38 39.2 51 10 US-09-864-761-42265 Sequence 4265, A

23 38 39.2 72 10 US-09-925-301-11562 Sequence 1562, Ap

24 38 39.2 89 10 US-09-815-242-11224 Sequence 11224, A

25 38 39.2 89 10 US-09-815-242-11245 Sequence 11245, A

26 38 39.2 305 10 US-09-815-242-12337 Sequence 12337, A

27 38 39.2 517 10 US-09-815-242-57232 Sequence 57232, Ap

28 38 39.2 517 10 US-09-815-242-57237 Sequence 57237, Ap

29 38 39.2 518 10 US-09-804-551b-8 Sequence 12650, A

30 38 39.2 532 9 US-09-738-626-6268 Sequence 6268, Ap

31 38 39.2 695 10 US-09-864-761-42967 Sequence 42967, A

32 38 39.2 2491 9 US-10-000-789-2 Sequence 2, Appl1

33 38 39.2 2548 10 US-09-851-682a-1 Sequence 1, Appl1

34 37.5 38.7 413 10 US-09-925-299-977 Sequence 977, App

35 37.5 38.7 1523 12 US-10-011-064-2 Sequence 2, Appl1

36 37.5 38.7 1523 12 US-10-052-586-290 Sequence 290, App

37 37 38.1 67 10 US-09-764-877-1827 Sequence 1827, Ap

38 37 38.1 188 9 US-09-808-602-109 Sequence 109, App

39 37 38.1 299 9 US-09-992-598-213 Sequence 213, App

40 37 38.1 299 9 US-09-989-293a-213 Sequence 213, App

41 37 38.1 299 9 US-09-989-735-213 Sequence 213, App

42 37 38.1 299 9 US-09-989-444-213 Sequence 213, App

43 37 38.1 299 9 US-09-989-436-213 Sequence 213, App

44 37 38.1 299 9 US-09-990-436-213 Sequence 213, App

45 37 38.1 299 9 US-09-991-181-213 Sequence 213, App

ALIGNMENTS

RESULT 1

US-09-834-795a-4

Sequence 4, Application US/09834795A

Patent No. US20020076710A1

GENERAL INFORMATION:

APPLICANT: Lawrence, Papsidero

APPLICANT: Lyn, Dyster

TITLE OF INVENTION: Detection and Treatment of Breast Cancer

CURRENT APPLICATION NUMBER: US/09/834, 795A

PRIOR FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 09/146, 580

PRIOR FILING DATE: 1998-09-03

PRIOR APPLICATION NUMBER: 60/071, 899

PRIOR FILING DATE: 1998-01-20

PRIOR APPLICATION NUMBER: 60/092, 155

PRIOR FILING DATE: 1998-07-09

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn version 3.0

SEQ ID NO 4

LENGTH: 16

TYPE: PRT

ORGANISM: Homo sapiens

US-09-834-795a-4

Query Match 100.0%; Score 97; DB 10; Length 16;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KNGKGNVCHRRKHGK 16

Db 1 KNGKGNVCHRRKHGK 16

RESULT 2

US-09-834-795a-2

Sequence 2, Application US/09834795A

Patent No. US20020076710A1

GENERAL INFORMATION:

APPLICANT: Lawrence, Papsidero

APPLICANT: Lyn, Dyster

APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (68)..(68)
OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
US-09-834-795A-2

Query Match 100.0%; Score 97; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNGKGNVCHRRKHHGK 16
|||||

DB 69 KNGKGNVCHRRKHHGK 84

RESULT 3

US-09-898-751A-6
Sequence 6, Application US/09898751A
Patent No. US20020160024A1
GENERAL INFORMATION:
APPLICANT: Oldham, Elizabeth R.
APPLICANT: Soto, Hortensia
APPLICANT: Liu, Ying
APPLICANT: Hudak, Susan A.
APPLICANT: Homey, Bernhard
APPLICANT: Morales, Janine M.
APPLICANT: Kellerman, Strid-Almee
APPLICANT: McEvoy, Leslie M.
APPLICANT: Bowman, Edward P.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
FILE REFERENCE: DX0882XK
CURRENT APPLICATION NUMBER: US/09/898,751A
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: US09/471,549
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: US60/136,570
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US60/113,858
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (529)..(529)
OTHER INFORMATION: unknown amino; may be "A", "C", or "G"
US-09-898-751A-6

Query Match 100.0%; Score 97; DB 9; Length 127;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KNGKGNVCHRRKHHGK 16
|||||

DB 92 KNGKGNVCHRRKHHGK 107

RESULT 4

US-09-813-492-2
Sequence 2, Application US/09813492
Patent No. US2002009735A1
GENERAL INFORMATION:
APPLICANT: Labow, Mark A.
APPLICANT: Mickanin, Craig Stephen
APPLICANT: Bhatia, Umesh
TITLE OF INVENTION: MAMMARY GLAND CHEMOKINE
FILE REFERENCE: 12345
CURRENT APPLICATION NUMBER: US/09/813,492
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 127
TYPE: PRT
ORGANISM: HUMAN
US-09-813-492-2

Query Match 100.0%; Score 97; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNGKGNVCHRRKHHGK 16
|||||

DB 92 KNGKGNVCHRRKHHGK 107

RESULT 5

US-09-834-795A-1
Sequence 1, Application US/09834795A
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (91)..(91)
OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-834-795A-1

Query Match 100.0%; Score 97; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
Db 92 KNGKGNVCHRRKHHGK 107

RESULT 6

US-09-931-381A-2
; Sequence 2, Application US/09931381A
; Patent No. US20020137107A1
; GENERAL INFORMATION:
; APPLICANT: Butcher, Eugene C.
; APPLICANT: Kunkel, Eric J.
; APPLICANT: Pan, Junliang
; APPLICANT: Soler-Ferran, Dulce
; TITLE OF INVENTION: Method for Identifying Agents Which
; TITLE OF INVENTION: Modulate Chemokine "Mec"-Induced Functions of CCR3 and/or
; TITLE OF INVENTION: CCR3
; FILE REFERENCE: 1855, 2010-003
; CURRENT APPLICATION NUMBER: US/09/931, 381A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: U.S. 09/638, 914
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(22)
US-09-931-381A-2

Query Match 100.0%; Score 97; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
Db 92 KNGKGNVCHRRKHHGK 107

RESULT 7

US-09-898-751A-10
; Sequence 10, Application US/09898751A
; Patent No. US20020160024A1
; GENERAL INFORMATION:
; APPLICANT: Oldham, Elizabeth R.
; APPLICANT: Solo, Hortensia
; APPLICANT: Liu, Ying
; APPLICANT: Hudak, Susan A.
; APPLICANT: Homey, Bernhard
; APPLICANT: Morales, Janine M.
; APPLICANT: Kellerman, Silda-Almee
; APPLICANT: McEvoy, Leslie M.
; APPLICANT: Bowman, Edward P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
; FILE REFERENCE: DX0882XK
; CURRENT APPLICATION NUMBER: US/09/898, 751A
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: US09/471,549
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US60/136,570
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US60/113,858
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 130
; TYPE: PRT

ORGANISM: Mus musculus
US-09-898-751A-10

Query Match 44.3%; Score 43; DB 9; Length 130;
Best Local Similarity 66.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRK 12
Db 92 KNGKGNVCHRRK 103

RESULT 8

US-09-950-933A-43
; Sequence 43, Application US/09950933A
; Patent No. US20020165141A1
; GENERAL INFORMATION:
; APPLICANT: Simmons, Carl R.
; APPLICANT: Navarro, Pedro
; TITLE OF INVENTION: Antimicrobial Peptides and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 35718/238472
; CURRENT APPLICATION NUMBER: US/09/950, 933A
; CURRENT FILING DATE: 2001-09-11
; PRIOR APPLICATION NUMBER: 60/232,569
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Zea mays
US-09-950-933A-43

Query Match 42.8%; Score 41.5; DB 9; Length 129;
Best Local Similarity 40.7%; Pred. No. 17;
Matches 11; Conservative 2; Mismatches 3; Indels 11; Gaps 2;

OY 1 KNGKGNV-----CHRRKHHGK 16
Db 60 KDGKGNLKPSCGCGECCRRRCSTHKK 86

RESULT 9

US-09-900-425A-2
; Sequence 2, Application US/09900425A
; Patent No. US20020164601A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Hongjiang
; APPLICANT: Crooke, Stanley T.
; TITLE OF INVENTION: Human RNase III and Compositions and Uses Thereof
; FILE REFERENCE: ISPH-0522
; CURRENT APPLICATION NUMBER: US/09/900, 425A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US 09/479, 783
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 08/870,608
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: US 80/659,440
; PRIOR FILING DATE: 1996-06-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-900-425A-2

Query Match 42.3%; Score 41; DB 9; Length 1374;
Best Local Similarity 54.5%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 4 KGNVCHRRKHH 14

DB 3 QGNTCHMSFH 13

RESULT 10

US-09-738-626-6282

Sequence 6282, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738,626

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 6282

LENGTH: 166

TYPE: PR

ORGANISM: Corynebacterium glutamicum

US-09-738-626-6282

Query Match 41.2%; Score 40; DB 9; Length 166;
Best Local Similarity 60.0%; Pred. No. 35;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 NVCHRRKHHG 15
DB 53 NUCGRTNHG 62

RESULT 11

US-09-864-761-42897

Sequence 42897, Application US/09864761

Patent No. US20020048765A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 42897

LENGTH: 492

TYPE: PR

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC007228.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1

OTHER INFORMATION: SWISSPROT HIT: P10078, EVALU0 0.00e+00

OTHER INFORMATION: EST_HUMAN HIT: AL042491.2, EVALU0 0.00e+00

OTHER INFORMATION: EST_HUMAN HIT: AL042491.2, EVALU0 0.00e+00

US-09-864-761-42897

Query Match 41.2%; Score 40; DB 10; Length 492;
Best Local Similarity 50.0%; Pred. No. 96;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 VCHRRKHHG 16
DB 445 ICHRRSHTE 454

RESULT 12

US-09-823-038A-52

Sequence 52, Application US/09823038A

Patent No. US20020058335A1

GENERAL INFORMATION:

APPLICANT: Strachan, Lorne

APPLICANT: Sleeman, Matthew

APPLICANT: Abernethy, Nevin

APPLICANT: Onrust, Rene

APPLICANT: Kumble, Anand

APPLICANT: Murlison, Greg

TITLE OF INVENTION: Compositions Isolated From Stromal Cells

FILE REFERENCE: 11000.1037c3

CURRENT APPLICATION NUMBER: US/09/823,038A

CURRENT FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 61

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 52

LENGTH: 757

TYPE: PR

ORGANISM: Mouse

US-09-823-038A-52

Query Match 41.2%; Score 40; DB 10; Length 757;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 5 GNGVCHRRKHHG 15
| | | | : | |
Db 130 GVGCHPRRQHG 140

RESULT 13
US-09-764-864-1435
; Sequence 1435, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1435
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1435

Query Match 40.7%; Score 39.5; DB 10; Length 313;
Best Local Similarity 33.3%; Pred. No. 75;
Matches 9; Conservative 1; Mismatches 2; Indels 15; Gaps 1;

QY 2 NGKGNVC-----HRRKH 13
| | | | : | | |
Db 239 NGKGVACFCPCICKTAALGMHQKH 265

RESULT 14
US-09-764-847-504
; Sequence 504, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 504
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-847-504

Query Match 40.2%; Score 39; DB 10; Length 83;
Best Local Similarity 40.0%; Pred. No. 25;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNGKGNVCHRRKHHG 15
: | | | : | | |
Db 45 RDGGGKYTRRQHG 59

RESULT 15
US-09-798-042-101
; Sequence 101, Application US/09798042
; Patent No. US20020068343A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS

; TITLE OF INVENTION: AND TREATMENT OF EHRILICHIA INFECTION
; FILE REFERENCE: 210121.439c7
; CURRENT APPLICATION NUMBER: US/09/798,042
; CURRENT FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 101
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Ehrlichia sp.
US-09-798-042-101

Query Match 40.2%; Score 39; DB 10; Length 147;
Best Local Similarity 58.3%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NGKGNVCHRRKH 13
| | | | : | | |
Db 43 NIDGVCRRRKH 54

Search completed: January 14, 2003, 18:27:29
Job time : 3.13333 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:11:25 ; Search time 3.55556 Seconds
(without alignments)
132.403 Million cell updates/sec

Title: US-09-834-794a-4
Perfect score: 97
Sequence: 1 KNKGNCVCHRRKHGK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	97	100.0	16 4 US-09-146-580-4	Sequence 4, Appli
2	97	100.0	104 4 US-09-146-580-2	Sequence 2, Appli
3	97	100.0	127 4 US-09-146-580-1	Sequence 1, Appli
4	41	42.3	346 4 US-09-049-672A-28	Sequence 28, Appli
5	40	41.2	583 1 US-08-445-586-9	Sequence 9, Appli
6	40	41.2	583 1 US-08-484-493-14	Sequence 14, Appli
7	40	41.2	583 1 US-08-484-494-14	Sequence 14, Appli
8	40	41.2	583 2 US-08-345-212-14	Sequence 14, Appli
9	40	41.2	583 4 US-09-249-003-14	Sequence 14, Appli
10	40	41.2	801 4 US-09-134-001C-5584	Sequence 5584, Ap
11	39	40.2	185 4 US-08-975-762-10	Sequence 10, Appli
12	39	40.2	185 4 US-08-821-324-10	Sequence 10, Appli
13	39	40.2	185 4 US-09-295-028-10	Sequence 10, Appli
14	39	40.2	185 4 US-09-106-582-10	Sequence 10, Appli
15	39	40.2	895 1 US-08-123-161A-8	Sequence 8, Appli
16	39	40.2	895 1 US-08-483-278-8	Sequence 8, Appli
17	38	39.2	260 4 US-09-174-768-6	Sequence 6, Appli
18	38	39.2	2546 4 US-09-172-422-1	Sequence 1, Appli
19	37.5	38.7	1523 4 US-09-182-024A-2	Sequence 2, Appli
20	37	38.1	38 1 US-08-287-717-9	Sequence 9, Appli
21	37	38.1	38 1 US-08-481-888A-1	Sequence 1, Appli
22	37	38.1	38 1 US-08-485-273A-1	Sequence 1, Appli
23	37	38.1	38 1 US-08-441-914-9	Sequence 9, Appli
24	37	38.1	38 2 US-08-973-563A-1	Sequence 1, Appli
25	37	38.1	38 2 US-08-973-559-1	Sequence 1, Appli
26	37	38.1	52 3 US-08-978-741-15	Sequence 15, Appli
27	37	38.1	52 4 US-09-333-729A-15	Sequence 15, Appli

28	37	38.1	312 1	US-08-247-908A-2	Sequence 2, Appli
29	37	38.1	312 1	US-08-453-942-2	Sequence 2, Appli
30	37	38.1	312 2	US-08-926-885A-2	Sequence 2, Appli
31	37	38.1	312 5	PCT-US94-05290-2	Sequence 2, Appli
32	37	38.1	424 1	US-08-247-908A-11	Sequence 11, Appli
33	37	38.1	424 1	US-08-453-942-11	Sequence 11, Appli
34	37	38.1	424 2	US-08-926-885A-11	Sequence 11, Appli
35	37	38.1	424 5	PCT-US94-05290-11	Sequence 11, Appli
36	37	38.1	1786 2	US-08-477-451-16	Sequence 16, Appli
37	36.5	37.6	714 4	US-09-115-954-4	Sequence 4, Appli
38	36.5	37.6	1032 4	US-09-115-954-8	Sequence 8, Appli
39	36.5	37.6	1044 4	US-09-115-954-2	Sequence 2, Appli
40	36	37.1	23 3	US-08-993-235-2	Sequence 2, Appli
41	36	37.1	24 1	US-08-481-888A-5	Sequence 5, Appli
42	36	37.1	24 1	US-08-485-273A-5	Sequence 5, Appli
43	36	37.1	24 2	US-08-973-563A-5	Sequence 5, Appli
44	36	37.1	24 2	US-08-973-559-5	Sequence 5, Appli
45	36	37.1	25 1	US-07-689-693B-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-146-580-4
Sequence 4, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papsidero, Lawrence D
APPLICANT: Dyster, Lyn M
APPLICANT: Frustraci, Jana M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
CURRENT FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 16
TYPE: PRT
ORGANISM: Homo sapiens
US-09-146-580-4
Query Match          100.0%; Score 97; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB      1 KNKGNCVCHRRKHGK 16
      1 KNKGNCVCHRRKHGK 16
RESULT 2
US-09-146-580-2
Sequence 2, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papsidero, Lawrence D
APPLICANT: Dyster, Lyn M
APPLICANT: Frustraci, Jana M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
CURRENT FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)
; OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
US-09-146-580-2

Query Match      100.0%; Score 97; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
Db 69 KNGKGNVCHRRKHHGK 84

RESULT 3
US-09-146-580-1
; Sequence 1, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyaster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (70)
; OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
; NAME/KEY: UNSURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-146-580-1

Query Match      100.0%; Score 97; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
Db 92 KNGKGNVCHRRKHHGK 107

RESULT 4
US-09-049-672A-28
; Sequence 28, Application US/09049672A
; Patent No. 6135941
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
```

```

; APPLICANT: Yue, Henry
; APPLICANT: Au-Yang, Janice
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: HUMAN IMMUNE SYSTEM ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,672A
; FILING DATE: HERewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0497 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 346 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1550785
US-09-049-672A-28

Query Match      42.3%; Score 41; DB 4; Length 346;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 5 GNVCHRRKHHGK 16
Db 104 GCACHOKLHFGQ 115

RESULT 5
US-08-445-586-9
; Sequence 9, Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Ito, Toshimi
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Amanu, Egon
; TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
```


RESULT 7
 US-08-484-494-14
 Sequence 14, Application US/08484494
 Patent No. 5798239
 GENERAL INFORMATION:
 APPLICANT: Wilson, Peter J
 APPLICANT: Morris, Charles P
 APPLICANT: Anson, Donald S
 APPLICANT: Occhiodoro, Teresa
 APPLICANT: Bielicki, Julie
 APPLICANT: Clements, Peter R
 APPLICANT: Hopwood, John J
 TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Scully, Scott, Murphy & Presser
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: New York
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,494
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 991,973
 FILING DATE: 17-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Digiglio, Frank S
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 84162
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:

LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-494-14

Query Match 41.2%; Score 40; DB 1; Length 583;
Best Local Similarity 42.9%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 NGKGNVCHRRKHG 15
Db 142 HSKTDFCHHPLHNG 155

RESULT 8
US-08-345-212-14

Sequence 14, Application US/08345212
Patent No. 5932211

GENERAL INFORMATION:

APPLICANT: Wilson, Peter J

APPLICANT: Morris, Charles P

APPLICANT: Anson, Donald S

APPLICANT: Occhiodoro, Teresa

APPLICANT: Bielicki, Julie

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

Db 142 HSKTDFCHHPLHNG 155

RESULT 9
US-09-249-003-14

Sequence 14, Application US/09249003
Patent No. 6153188

GENERAL INFORMATION:

APPLICANT: Wilson, Peter J

APPLICANT: Morris, Charles P

APPLICANT: Anson, Donald S

APPLICANT: Occhiodoro, Teresa

APPLICANT: Bielicki, Julie

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

APPLICANT: Hopwood, John J

APPLICANT: Clements, Peter R

US-09-249-003-14

Query Match 41.2%; Score 40; DB 4; Length 583;
Best Local Similarity 42.9%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 NGKGNVCHRRKHG 15
Db 142 HSKTDFCHHPLHNG 155

RESULT 10
US-09-134-001C-5584

Sequence 5584, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al

APPLICANT: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC

APPLICANT: EPIDEMIOLOGICAL AND THERAPEUTICS

APPLICANT: GTC-007

APPLICANT: US/09134, 001C

APPLICANT: US/09134, 001C

APPLICANT: US/09134, 001C

APPLICANT: US/09134, 001C

;; CURRENT FILING DATE: 1998-08-13
;; PRIOR APPLICATION NUMBER: US 60/064,964
;; PRIOR FILING DATE: 1997-11-08
;; PRIOR APPLICATION NUMBER: US 60/055,779
;; PRIOR FILING DATE: 1997-08-14
;; NUMBER OF SEQ ID NOS: 5674
;; SEQ ID NO 5584
;; LENGTH: 801
;; TYPE: PRT
;; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5584

Query Match 41.2%; Score 40; DB 4; Length 801;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 NGKGNVCHRRKH 16
DB 755 RKGKGGKKKKRKGK 770

RESULT 11
US-08-975-762-10
Sequence 10, Application US/08975762
Patent No. 6207169
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,762
FILING DATE: 21-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Ehrlichia
US-08-975-762-10

Query Match 40.2%; Score 39; DB 4; Length 185;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 2 NGKGNVCHRRKH 13
DB 43 NIDGKVCRRKH 54

RESULT 12
US-08-821-324-10
Sequence 10, Application US/08821324
Patent No. 6231869
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Ehrlichia
US-08-821-324-10

Query Match 40.2%; Score 39; DB 4; Length 185;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 NGKGNVCHRRKH 13
DB 43 NIDGKVCRRKH 54

RESULT 13
US-09-295-028-10
Sequence 10, Application US/09295028
Patent No. 6277381
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
AND TREATMENT OF EHRLICHIA INFECTION
FILE REFERENCE: 210121.439C4
CURRENT APPLICATION NUMBER: US/09/295,028
CURRENT FILING DATE: 1999-04-20
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 185
TYPE: PRT

ORGANISM: Ehrlichia sp.
US-09-293-028-10

Query Match 40.2%; Score 39; DB 4; Length 185;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NGKGNVCHRRKH 13
DB 43 NIDGKVCRRKH 54

RESULT 14

US-09-106-582-10
Sequence 10, Application US/09106582
Patent No. 6306402
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/106,582
FILING DATE: 29-JUN-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Ehrlichia
US-09-106-582-10

Query Match 40.2%; Score 39; DB 4; Length 185;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NGKGNVCHRRKH 13
DB 43 NIDGKVCRRKH 54

RESULT 15

US-08-123-161A-8
Sequence 8, Application US/08123161A
Patent No. 5449616
GENERAL INFORMATION:
APPLICANT: Campbell, Kevin P.
APPLICANT: Roberts, Steven L.
APPLICANT: Anderson, Richard D.

APPLICANT: Ibraghimov, Oxana B.
APPLICANT: Yang, Bin
TITLE OF INVENTION: NUCLEIC ACID ENCODING DYSTROPHIN-ASSOCIATED
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: ME
COUNTRY: USA
ZIP: 03911

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,161A
FILING DATE: 16-SEP-93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/946,234
FILING DATE: 14-SEP-92
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M.
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: UIRF89-11A4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (207) 363-0558
TELEFAX: (207) 363-0528
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 895 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-123-161A-8

Query Match 40.2%; Score 39; DB 1; Length 895;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 VCHRRKHRRKH 16
DB 773 ICYKRRKHRRKH 782

Search completed: January 14, 2003, 18:17:05
Job time: 4.55556 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:42:25 ; Search time 2.6 Seconds

(without alignments)
287.144 Million cell updates/sec

Title: US-09-834-794A-3

Perfect score: 94

Sequence: 1 TEVSHHSRRLRRVNC 18

Scoring table: BLOSUM62

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt,40:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	127	1	SY28_HUMAN
2	75	79.8	130	1	SY28_MOUSE
3	49	52.1	400	1	VG78_HSV1
4	48	51.1	1246	1	SKIW_HUMAN
5	44.5	47.3	241	1	PDGB_SHEEP
6	42	44.7	271	1	DLIA_ECOLI
7	42	44.7	1108	1	MAN2_DROME
8	41	43.6	700	1	NAD2_SCHPO
9	41	43.6	1124	1	PHYA_LATSA
10	41	43.6	1124	1	PHYA_PEA
11	40.5	43.1	558	1	VP10_RBSDV
12	40	42.6	158	1	ML43_ARATH
13	40	42.6	222	1	MYLJ_PSEPU
14	40	42.6	429	1	MYLJ_RUEGE
15	39.5	42.0	226	1	TSIS_SMSAV
16	39.5	42.0	241	1	PDGB_HUMAN
17	39	41.5	201	1	HIS5_THEMA
18	39	41.5	4385	1	VP73_CAEEL
19	38.5	41.0	245	1	PDGB_CAEEL
20	38	40.4	116	1	PDGB_FELCA
21	38	40.4	116	1	SP21_BACST
22	38	40.4	238	1	V554_AOQAE
23	38	40.4	328	1	STRE_STRGR
24	38	40.4	450	1	RUMA_RALSO
25	38	40.4	638	1	SCAD_HUMAN
26	38	40.4	638	1	SCAD_PANTR
27	38	40.4	1122	1	VG3C_YEAST
28	38	40.4	1375	1	BNRI_YEAST
29	37.5	39.9	248	1	AKAC_HUMAN
30	37	39.4	250	1	PIGL_SCHPO
31	37	39.4	266	1	UNG_HCMVA
32	37	39.4	324	1	VYWK_CAEEL
33	37	39.4	334	1	Y700_METUA
					YCL2_ECOLI
					004871 escherichia

34	37	39.4	365	1	Y113_YEAST
35	37	39.4	389	1	Y466_TREPA
36	37	39.4	399	1	RPP_P14HB
37	37	39.4	653	1	PABP_SCHPO
38	37	39.4	896	1	DSC3_HUMAN
39	37	39.4	922	1	W70T_MOUSE
40	36	38.3	78	1	Y386_HELPY
41	36	38.3	116	1	Y1B_SOCMV
42	36	38.3	165	1	Y4UD_RHISN
43	36	38.3	166	1	Y4A3_HAEIN
44	36	38.3	177	1	NIUG_PSEAE
45	36	38.3	231	1	ELI5_HORVU

ALIGNMENTS

RESULT 1
SY28_HUMAN STANDARD: PRT; 127 AA.
AC QGNRJ3;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-OCT-2001 (Rel. 41, Last annotation update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Small inducible cytokine A28 precursor (CCL28) (Mucosae-associated epithelial chemokine) (MEC) (CCK1 protein).
GN SCYA28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RC TISSUE=Fetal heart, and Osteoblast;
RX MEDLINE=20357357; PubMed=10781587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Homey B., Catron D., Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J., Kershenovich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.;
RT "Identification of a novel chemokine (CCL28), which binds CCR10 (GPR2).";
RT J. Biol. Chem. 275:22313-22323(2000).
RL T.
RM
RN
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RX MEDLINE=20432268; PubMed=10975800;
RA Pan J., Kunkel E.J., Goesslar U., Lazarus N., Langdon P., Broadwell K., Vieira M.A., Genovese M.C., Butcher E.C., Soler D.;
RT "A novel chemokine ligand for CCR10 and CCR3 expressed by epithelial cells in mucosal tissues";
RT J. Immunol. 165:2943-2949(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Zhang W., He L., Yuan Z., Wan T., Cao X.;
RT "A novel CC chemokine homology with TECK";
RT Submitted (Dec-1998) to the EMBL/Genbank/DBJ databases.
RL
RM
RN
RP FUNCTION: CHEMOKINE ACTIVITY FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS. BINDS TO CCR3 AND CCR10 AND INDUCES CALCIUM MOBILIZATION IN A DOSE-DEPENDENT MANNER.
RL
RM
RN
RP TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED BY EPITHELIAL CELLS OF DIVERSE TISSUES INCLUDING NORMAL AND PATHOLOGICAL COLON, SALIVARY GLAND, MAMMARY GLAND, TRACHEA AND RECTUM. ALSO FOUND IN PROSTATE, SPLEEN, THYROID, PSORIASIS SKIN AND IN LOWER LEVELS IN PERIPHERAL BLOOD LEUKOCYTES, SMALL INTESTINE, PEYER'S PATCHES, STOMACH AND NORMAL SKIN.
RL
RM
RN
RP SIMILARITY: BELONGS TO THE INTERFERON BETA FAMILY (SMALL CYTOKINE C-C) (CHEMOKINE CC).
RL
RM
RN
RP This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@sb-sib.ch.

CC EMBL: AF220210; AAF87205.1; -

DR EMBL: AF266504; AAG16691.1; -

DR EMBL: AF110384; AAG43193.1; -

DR MIM: 605240; -

DR InterPro: IPR000827; CC-Chemkine_sml.

DR InterPro: IPR001811; Chemokine_IL8.

DR SMART: SM00199; SCY: 1.

DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.

KW Cytokine; Chemotaxis; Signal.

FT SIGNAL 1 19

FT CHAIN 20 127

FT DISULFID 30 58

FT DISULFID 31 73

FT CARBOHYD 78 78

FT SEQUENCE 127 AA; 14280 MW; 3E8551A63A2C8D62 CRC64;

Query Match 100.0%; Score 94; DB 1; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLRRVNC 18
DB 32 TEVSHHSRRLRRVNC 49

RESULT 2
SY28_MOUSE
ID SY28_MOUSE STANDARD; PRT; 130 AA.
AC O9JIL2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Small inducible cytokine A28 precursor (CCL28).
GN SCYA28.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
[1]
RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
RC TISSUE=Kidney.
RX MEDLINE=20357357; PubMed=10781587;
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Homey B., Catron D.,
Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J.,
Kershenovich D., Smith K., McClanahan T., Vicari A.P., Zlotnik A.;
RT Identification of a novel chemokine (CCL28), which binds CCR10
(GPR2).
RL J. Biol. Chem. 275:22313-22323(2000).
CC -1- FUNCTION: CHEMOTACTIC FOR RESTING CD4, CD8 T-CELLS AND EOSINOPHILS
(BY SIMILARITY). BINDS TO CCR10 AND INDUCES CALCIUM MOBILIZATION
IN A DOSE-DEPENDENT MANNER.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC OF NORMAL COLON, KIDNEY, PEYER'S PATCHES, LYMPH NODES. ALSO FOUND
IN LOWER LEVELS IN BRAIN, SPLEEN AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE
C-C) (CHEMOKINE CC).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sb-sib.ch).

CC EMBL: AF220238; AAF87206.1; -

DR MGD: MGI:161731; SCYA28.

DR InterPro: IPR000827; CC-Chemkine_sml.

DR InterPro: IPR001811; Chemokine_IL8.

DR SMART: SM00199; SCY: 1.

DR PROSITE: PS00472; SMALL_CYTOKINES_CC; FALSE_NEG.

KW Cytokine; Chemotaxis; Signal.

FT SIGNAL 1 16

FT CHAIN 17 130

FT DISULFID 30 58

FT DISULFID 31 73

FT CARBOHYD 78 78

FT SEQUENCE 130 AA; 14570 MW; 6F3B909A4F97F013 CRC64;

Query Match 79.8%; Score 75; DB 1; Length 130;
Best Local Similarity 77.8%; Pred. No. 8.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLRRVNC 18
DB 32 TEVSHHSRRLRRVSSC 49

RESULT 3
VG78_HSV11
ID VG78_HSV11 STANDARD; PRT; 400 AA.
AC Q00167;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical gene 78 zinc-binding protein.
GN 78.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCB1_TaxID=10401;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT Channel catfish virus: a new type of herpesvirus.*;
RL Virology 186:9-14(1992).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sb-sib.ch).

CC EMBL: M75136; AAB8180.1; -

DR PIR: D36794; ZBBB14.

KW Hypothetical protein; Zinc; Zinc-finger.

SQ SEQUENCE 400 AA; 44167 MW; D941DB7738B95CA9 CRC64;

Query Match 52.1%; Score 49; DB 1; Length 400;
Best Local Similarity 52.9%; Pred. No. 0.74;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 EVSHHSRRLRRVNC 18
DB 190 KVOHRHSRLLRRRC 206

RESULT 4
SK1W_HUMAN
ID SK1W_HUMAN STANDARD; PRT; 1246 AA.
AC O15477; O12902; O15005; O15476;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Helicase SK12W (Helicase-like protein) (HLP).
GN SK12L OR SK12 OR SK12W OR DDX13 OR W.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MDLINE=95334363; PubMed=7610041;
 RA Dangel A.W., Shen L., Mendoza A.R., Wu L.-C., Yu C.Y.;
 RT "Human helicase gene SKI2W in the HLA class III region exhibits
 striking structural similarities to the yeast antiviral gene SKI2 and
 RT to the human gene KIA0052: emergence of a new gene family.";
 RN Nucleic Acids Res. 23:2120-2126(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Bowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
 RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
 RT "Sequence determination of 300 kilobases of the human class III MHC
 RT locus.";
 RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Albertella M.R., Jones H., Thomson W., Olavesen M.G.,
 RA Campbell R.D.;
 RN Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE OF 1-958 FROM N.A.
 RC TISSUE=Hippocampus;
 RA Lee S.-G., Song K.;
 RT "Cloning of the human helicase-like protein gene homologous to yeast
 RT SKI2.";
 RN Korean J. Biochem. 26:215-219(1994).
 RN [5]
 RN SEQUENCE OF 230-960 FROM N.A.
 RC TISSUE=Liver;
 RX MDLINE=95278935; PubMed=7759100;
 RA Lee S.-G., Lee I., Park S.H., Kang C., Song K.;
 RT "Identification and characterization of a human cDNA homologous to
 RT yeast SKI2.";
 RN Genomics 25:660-666(1995).
 RN [6]
 RN REVISIONS.
 RA Song K.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: HELICASE; HAS ATPASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Nucleus (Potential).
 CC -1- SIMILARITY: BELONGS TO THE HELICASE FAMILY. SKI2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Z48796; CAAB8733.1; ALT_INT.
 DR EMBL: AF019413; AAB67978.1; -.
 DR EMBL: X98378; CAAB7024.1; -.
 DR EMBL: U09877; AAB52523.1; ALT_INT.
 DR Genew; HGNC:10898; SKI2L.
 DR MIM: 600478; -.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC_C; 1.
 KM Helicase; ATP-binding; RNA-binding.
 FT NP_BIND 332 339
 FT SITE 423 426
 FT STATE 66 120
 FT CONFLICT 151 151
 FT CONFLICT 253 253
 R -> Q (IN REF. 3, 4 AND 6).
 R -> Q (IN REF. 4).

FT CONFLICT 265 265 A -> P (IN REF. 4).
 FT CONFLICT 289 289 A -> P (IN REF. 4).
 FT CONFLICT 366 366 S -> T (IN REF. 4 AND 5).
 FT CONFLICT 562 583 RTAOLPVVTFVTFGRGDEDA ->
 FT CONFLICT 623 623 PHYSCPMKCPSPGCAVMSST (IN REF. 4).
 FT CONFLICT 914 914 H -> Q (IN REF. 1).
 FT CONFLICT 917 917 MISSING (IN REF. 4).
 FT CONFLICT 947 960 V -> M (IN REF. 2).
 FT CONFLICT 1052 1052 ROOPFKKDPPLAA -> SSSONSRRILPQP (IN
 FT CONFLICT 1071 1071 REF. 4).
 FT CONFLICT 1246 AA; 137799 MM; 904405C8C70621A0 CRC64;
 SO SEQUENCE F -> L (IN REF. 2, 3 AND 6).
 A -> V (IN REF. 2).
 Query Match 51.1%; Score 48; DB 1; Length 1246;
 Best Local Similarity 56.2%; Pred. No. 3.7;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 TEVSHHSRRLRLRN 16
 DB 819 TETQMIGRIMESYN 834
 RESULT 5
 PDGB_SHEEP STANDARD; PRT; 241 AA.
 ID AC 095229;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Platelet-derived growth factor, B chain precursor (PDGF B-chain)
 DE (PDGF-2).
 GN PDGB.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=TEXEL; TISSUE=Spleen;
 RA Woodall C.J., Zhang Z., Walt N.J.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
 CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
 CC APARENT RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
 CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
 CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
 CC -1- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
 CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
 CC TRANSFORMATION PROCESSES.
 CC -1- MISCELLANEOUS: A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE
 CC PDGF RECEPTOR.
 CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X97123; CAAB5790.1; -.
 DR HSPB; P01127; 1PDG.
 DR InterPro: IPR002400; GF_cysknot.
 DR InterPro: IPR000072; PD_growth_factor.
 DR Pfam: PF00341; PDGF; 1.
 DR PRINTS: PR00438; GFCSKNOT.
 DR PRODOM: PD001629; PD_growth_factor; 1.
 DR SMART: SM00141; PDGF; 1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS00278; PDGF_2; 1.

KW Mitogen: Growth factor; Proto-oncogene; Platelet; Signal.
 FT SIGMNL 1 20 BY SIMILARITY.
 FT PROPEP 21 81 BY SIMILARITY.
 FT CHAIN 82 190 PLATELET-DERIVED GROWTH FACTOR, B CHAIN.
 FT PROPEP 191 241 BY SIMILARITY.
 FT SITE 108 108 INVOLVED IN RECEPTOR BINDING.
 FT SITE 111 111 INVOLVED IN RECEPTOR BINDING.
 FT DISULFID 97 141 BY SIMILARITY.
 FT DISULFID 130 178 BY SIMILARITY.
 FT DISULFID 134 180 BY SIMILARITY.
 FT DISULFID 124 124 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 133 133 INTERCHAIN (BY SIMILARITY).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 241 AA: 27331 MW: 3781ECL12E7D2863 CRC64;
 Query Match 47.3%; Score 44.5; DB 1; Length 241;
 Best Local Similarity 68.8%; Pred. No. 2.5;
 Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
 QY 1 TEVSHHSRLRLRVN 16
 DB 101 TEVSHHSRLRLRVN 115
 RESULT 6
 ID DJLA_ECOLI STANDARD; PRT; 271 AA.
 AC P31680;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DnaJ-like protein dJ1A.
 GN DJLA OR B0055.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=9234977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RA Gregory J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mao B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RT Science 277:1453-1474(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mao B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RT Science 277:1453-1474(1997).
 RL [3]
 RN CHARACTERIZATION, AND SEQUENCE OF 1-45 FROM N.A.
 RP MEDLINE=96405649; PubMed=8809778;
 RA Clarke D.J., Jacq A., Holland I.B.;
 RA "A novel DnaJ-like protein in Escherichia coli inserts into the
 RT cytoplasmic membrane with a type III topology.";
 RL Mol. Microbiol. 20:1273-1286(1996).
 RL [4]
 RN CHARACTERIZATION, AND MUTAGENESIS OF HIS-233.
 RP MEDLINE=98030199; PubMed=9364917;
 RA Kelley W.L., Georgopoulos C.;
 RA "Positive control of the two-component RscC/B signal transduction
 RT network by DjaA: a member of the DnaJ family of molecular chaperones
 RL in Escherichia coli.";
 RL Mol. Microbiol. 25:913-931(1997).
 RL [5]
 RN MUTAGENESIS.

RX MEDLINE=98030200; PubMed=9364918;
 RA Clarke D.J., Holland I.B., Jacq A.;
 RT "Point mutations in the transmembrane domain of DjaA, a
 RT membrane-linked DnaJ-like protein, abolish its function in promoting
 RT colanic acid production via the Rcs signal transduction pathway.";
 RL Mol. Microbiol. 25:933-944(1997).
 CC -1- FUNCTION: CHAPERONE THAT MAY PLAY A ROLE IN THE CORRECT ASSEMBLY,
 CC ACTIVITY AND/OR MAINTENANCE OF A NUMBER OF MEMBRANE PROTEINS,
 CC INCLUDING TWO-COMPONENT SIGNAL-TRANSDUCTION SYSTEMS. MIGHT CO-
 CC OPERATE WITH DnaK TO ACTIVATE THE RCS TWO-COMPONENT SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Type III membrane protein.
 CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: D10483; -; NOT_ANNOTATED_CDS.
 DR EMBL: AEO00116; AAC73166.1; -
 DR Ecocore: EGI1570; dJ1A.
 DR InterPro: IPR001623; DnaJ_N.
 DR Pfam: PF00226; DnaJ; 1.
 DR SMART: SM00271; DnaJ; 1.
 DR PROSITE: PS00636; DnaJ_1; FALSE_NEG.
 DR PROSITE: PS50076; DnaJ_2; 1.
 KW Chaperone; Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 1 6 PERIPLASMIC (POTENTIAL).
 FT TRANSMEM 7 27 POTENTIAL.
 FT DOMAIN 28 271 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 205 271 J-DOMAIN.
 FT MUTAGEN 15 15 L->R: LOSS OF ACTIVATION OF RCS.
 FT MUTAGEN 16 16 M->R: ONLY PARTIAL ACTIVATION OF RCS.
 FT MUTAGEN 233 233 H->Q: LOSS OF ACTIVITY.
 SQ SEQUENCE 271 AA: 30579 MW: 80A0FC28F6D470DF CRC64;
 Query Match 44.7%; Score 42; DB 1; Length 271;
 Best Local Similarity 41.2%; Pred. No. 7.4;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TEVSHHSRLRLRVN 17
 DB 73 TEADHIAQLMDRMNL 89
 RESULT 7
 ID MAN2_DROME STANDARD; PRT; 1108 AA.
 AC Q24451;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-mannosidase II (EC 3.2.1.114) (Mannosyl-oligosaccharide
 DE 1,3-1,6-alpha-mannosidase) (MAN II) (Golgi alpha-mannosidase II)
 DE (AMAN II).
 GN ALPHA-MAN-II OR GMII.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95196999; PubMed=7890162;
 RA Foster J.M., Yuckin B., Lockyer A.E., Roberts D.B.;
 RA "Cloning and sequence analysis of Gmit, a Drosophila melanogaster
 RT homologue of the cDNA encoding murine Golgi alpha-mannosidase II.";
 RL Gene 154:183-186(1995).
 CC -1- FUNCTION: CATALYZES THE FIRST COMMITTED STEP IN THE BIOSYNTHESIS
 CC OF COMPLEX N-GLYCANS. IT CONTROLS CONVERTED STEP OF HIGH MANNOSE TO

CC COMPLEX N-GLYCANS: THE FINAL HYDROLYTIC STEP IN THE N-GLYCAN
 CC MATURATION PATHWAY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the terminal 1,3- and 1,6-linked
 CC alpha-D-mannose residues in the mannosyl-oligosaccharide
 CC Man(5)(GlcNAc)(3).
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 38 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X77652; CAA54732.1; -
 CC FLYBase: F89n0011740; Alpha-Man-II.
 CC InterPro: IPR000602; Glyco_hydro_38.
 CC Pfam: PF01074; Glyco_hydro_38: 1.
 CC KEGG: Glycosylase; Transmembrane; Signal-anchor: Golgi stack.
 CC FT DOMAIN 1
 CC TRANSMEM 10 30
 CC CYTOPLASMIC (POTENTIAL).
 CC (POTENTIAL).
 CC LUMENAL (POTENTIAL).
 CC DOMAIN 31 1108
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC S0 SEQUENCE 1108 AA; 126719 MW; 5DC157FA211AE12 CRC64;
 CC -----
 CC Query Match 44.7%; Score 42; DB 1; Length 1108;
 CC Best Local Similarity 61.5%; Pred. No. 34;
 CC Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 CC -----
 CC Oy 6 HISRLRLRYVMNC 18
 CC ||| :||: ||| |
 CC Db 953 HIRLRLRYVMNC 965

RESULT 8
 NADE_SCHPO STANDARD; PRT; 700 AA.
 ID MADE_SCHPO
 AC 074940;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Putative glutamine-dependent NAD(+) synthetase (EC 6.3.5.1) (NAD(+)
 DE synthase [glutamine-hydrolyzing]).
 GN SPCC553.02.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC NCBI_TaxID=4896;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=972;
 CC MEDLINE=21848401; PubMed=11859360;
 CC Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 CC Spouris J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 CC Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 CC Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 CC Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
 CC Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 CC James K., Jones L., Jones M., Leather S., McDonald D., McLean J.,
 CC Mooney P., Moule S., Mungall K., Murphy L., Nibbelc D., Ogell C.,
 CC Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 CC Raftery K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC Skellon J., Simmonds M., Squares R., Stevens K.,
 CC Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 CC Weltsch I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC Gabel C., Fuchs M., Filtz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaive V., Motier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt Z., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe."
 RA Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + deamido-NAD(+) + NH(3) + H(2)O = AMP +
 CC diphosphate + NAD(+).
 CC -1- PATHWAY: NAD biosynthesis.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE NAD
 CC SYNTHETASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL023704; CAA19255.1; -
 CC InterPro: IPR003694; NAD synthase.
 CC InterPro: IPR003010; Ntlase/CNhydase.
 CC Pfam: PF00795; CN_hydrolase; 1.
 CC Pfam: PF02540; NAD_synthase; 1.
 CC DR TIGRFAMS: TIGR00552; nade; 1.
 CC KW Hypothetical protein; Ligase; NAD: ATP-binding.
 CC FT DOMAIN 327 700
 CC NP_BIND 357 364
 CC FT ACT_SITE 359 359
 CC BY SIMILARITY.
 CC S0 SEQUENCE 700 AA; 79520 MW; F915943317F62FED CRC64;
 CC -----
 CC Query Match 43.6%; Score 41; DB 1; Length 700;
 CC Best Local Similarity 57.1%; Pred. No. 31;
 CC Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 CC -----
 CC Oy 4 SHHSRLRLRYVM 17
 CC ||| :||: ||| |
 CC Db 202 SHHRLRLRYVM 215

RESULT 9
 PHYA_LATSA STANDARD; PRT; 1124 AA.
 ID PHYA_LATSA
 AC P93673;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phytochrome type A.
 GN PHYA.
 OS Lathyrus sativus (Chickling vetch).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eustosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Lathyrus.
 CC NCBI_TaxID=3860;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=CV. C-24;
 CC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
 CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
 CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
 CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
 CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
 CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE
 CC RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
 CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RIBULOSE-

CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
 CC PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS
 CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
 CC -1- SIMILARITY: BELONGS TO THE PHYCOCHROME FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARRNT-SIM) DIMERIZATION DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U84970; AABA7994.1; -
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR004359; HIS_KIN_s1g.
 DR InterPro: IPR003661; HIS_KinA.
 DR InterPro: IPR000014; PAS_domain.
 DR InterPro: IPR001294; Phytochrome.
 DR Pfam: PF00360; phytochrome; 1.
 DR Pfam: PF00512; signal; 1.
 DR Pfam: PF00989; PAS; 2.
 DR Pfam: PF01590; GAF; 1.
 DR Pfam: PF02518; HATPase_c; 1.
 DR PRINTS: PRO1033; PHYTOCHROME.
 DR SMART: SM00065; GAF; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00388; HisKA; 1.
 DR SMART: SM00091; PAS; 2.
 DR TIGRFA: TIGR00229; sensory_box; 1.
 DR PROSITE: PS50109; HIS_KIN; 1.
 DR PROSITE: PS50112; PAS; 2.
 DR PROSITE: PS50045; PHYTOCHROME_1; 1.
 DR PROSITE: PS50046; PHYTOCHROME_2; 1.
 DR PROSITE: PS50046; PHYTOCHROME_2; 1.
 DR Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
 DR Repeat; Multigene family.
 FT DOMAIN 617 687 PAS 1.
 FT DOMAIN 750 821 PAS 2.
 FT DOMAIN 901 1120 HISTIDINE KINASE.
 FT BINDING 323 323 CHROMOPHORE (BY SIMILARITY).
 SO SEQUENCE 1124 AA; 124250 MW; C7A6955BAAC28BA CRC64;
 Query Match 43.6%; Score 41; DB 1; Length 1124;
 Best Local Similarity 57.1%; Pred. No. 52;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VSHHISRLRLERNV 16
 DB 1083 ISLHISRLKLKLMN 1096
 RESULT 10
 ID PHYA_PEA STANDARD; PRT; 1124 AA.
 AC P15001;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phytochrome A.
 GN PHYA.
 OS Pisum sativum (Garden pea).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids 1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Alaska;

RA Sato N.;
 RT "Nucleotide sequence and expression of the phytochrome gene in Pisum
 RL sativum: differential regulation by light of multiple transcripts.";
 RL Plant Mol. Biol. 11:697-710(1988).
 CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
 CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT. THE PR FORM THAT ABSORBS
 CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PER FORM THAT
 CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
 CC PER INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
 CC RECONVERSION OF PER TO PR CANCELS THE INDUCTION OF THOSE
 CC RESPONSES. PER CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
 CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RUBULOSE-
 CC BISPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
 CC PROTOCHLOROPHYLLIDE REDUCTASE, RNA, ETC. IT ALSO CONTROLS
 CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
 CC -1- SIMILARITY: BELONGS TO THE PHYCOCHROME FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARRNT-SIM) DIMERIZATION DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M37217; AAA33682.1; -
 DR EMBL: X14077; CAA32242.1; -
 DR PIR: S06856; S06856.
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR004359; HIS_KIN_s1g.
 DR InterPro: IPR003661; HIS_KinA.
 DR InterPro: IPR000014; PAS_domain.
 DR InterPro: IPR001294; Phytochrome.
 DR Pfam: PF00360; phytochrome; 1.
 DR Pfam: PF00512; signal; 1.
 DR Pfam: PF00989; PAS; 2.
 DR Pfam: PF01590; GAF; 1.
 DR Pfam: PF02518; HATPase_c; 1.
 DR PRINTS: PRO1033; PHYTOCHROME.
 DR SMART: SM00065; GAF; 1.
 DR SMART: SM00387; HATPase_c; 1.
 DR SMART: SM00388; HisKA; 1.
 DR SMART: SM00091; PAS; 2.
 DR TIGRFA: TIGR00229; sensory_box; 1.
 DR PROSITE: PS50109; HIS_KIN; 1.
 DR PROSITE: PS50112; PAS; 2.
 DR PROSITE: PS50045; PHYTOCHROME_1; 1.
 DR PROSITE: PS50046; PHYTOCHROME_2; 1.
 DR PROSITE: PS50046; PHYTOCHROME_2; 1.
 DR Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
 DR Repeat; Multigene family.
 FT DOMAIN 617 687 PAS 1.
 FT DOMAIN 750 821 PAS 2.
 FT DOMAIN 901 1120 HISTIDINE KINASE.
 FT BINDING 323 323 CHROMOPHORE.
 SO SEQUENCE 1124 AA; 124027 MW; D94AD89F6AA72BD CRC64;
 Query Match 43.6%; Score 41; DB 1; Length 1124;
 Best Local Similarity 57.1%; Pred. No. 52;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 3 VSHHISRLRLERNV 16
 DB 1083 ISLHISRLKLKLMN 1096
 RESULT 11
 ID VP10_RBSDV STANDARD; PRT; 558 AA.
 VP10_RBSDV

```

AC P19898:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Protein S10.
GN S10.
OS Rice black streaked dwarf virus (RBSDV).
OC Viruses; dsRNA viruses; Reoviridae; Filoviridae.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RA Uyeda I., Azuhata F., Shikata E.
RT "Nucleotide sequence of rice black-streaked dwarf virus genome
segment 10."
RL Proc. Jpn. Acad., B, Phys. Biol. Sci. 66:37-40(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00606; BAA00481.1; -
DR PIR: J00409; MMXRT.
SQ SEQUENCE 558 AA; 63266 MW; 12D4C6D5959E3DC3 CRC64;

Query Match 43.1%; Score 40.5; DB 1; Length 558;
Best Local Similarity 36.4%; Pred. No. 29;
Matches 8; Conservative 7; Mismatches 2; Indels 5; Gaps 1;

QY 1 TEVSHH-----ISRRLERVNM 17
   |||||  |||||
Db 229 TDVTHYGGYDQFSRQMFERLNL 250

RESULT 12
ML43_ARATH STANDARD; PRT; 158 AA.
ID ML43_ARATH
AC 09SKS5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE MLP-Like protein 43.
GN MLP43 OR AT1G70890 OR F15H11.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV, Columbia.
RA Muller S., Klime S., Hauser M.T.;
RT "Molecular and phylogenetic analysis of a gene family in Arabidopsis
RT thaliana with similarities to major latex, pathogenesis-related and
RT ripening-induced proteins."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

[2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Columbia;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,

```

```

RA Miltcher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, Columbia.
RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC)."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ306140; CAC83578.1; -
DR EMBL: AC008148; AAD55504.1; -
DR EMBL: AY065135; AAL38311.1; -
KW Multigene family.
SQ SEQUENCE 158 AA; 17889 MW; 44ADA2C5C6C4C5A CRC64;

Query Match 42.6%; Score 40; DB 1; Length 158;
Best Local Similarity 33.3%; Pred. No. 9;
Matches 6; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 TEVSHHISRRLERVNM 18
   |||||  |||||
Db 29 TERPHVSKATPKDKIHGC 46

RESULT 13
XYLJ_PSEPU STANDARD; PRT; 222 AA.
ID XYLJ_PSEPU
AC 01NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 2-Hydroxypent-2,4-dienate hydratase (EC 4.2.-.-) (HPH) (2-oxopent-4-
DE enate hydratase).
GN XYLJ.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-22.
RC MEDLINE=92167811; PubMed=1791759;
RA Horn J.M., Harayama S., Timmis K.N.;
RT "DNA sequence determination of the TOL plasmid (pMWO) xylGf genes of
RT Pseudomonas putida: implications for the evolution of aromatic
RT catabolism."
RL Mol. Microbiol. 5:2459-2474(1991).
CC -----
CC -1- FUNCTION: CONVERSION OF 2-HYDROXPENT-2,4-DIENATE INTO 4-
CC HYDROXY-2-OXOPENTANOATE.
CC -1- PATHWAY: META-CLEAVAGE PATHWAY FOR THE DEGRADATION OF PHENOLS,
CC CRESOLS AND CATECHOL. PHENOL METABOLISM.
CC -1- SIMILARITY: BELONGS TO THE TOLU/XYL/XYLJ/HPG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 14, 2003, 18:15:35 ; Search time 2.4 Seconds
(without alignments)
145.508 Million cell updates/sec

Title: US-09-834-794a-3
Perfect score: 94
Sequence: 1 TEVSHHSRRLRLRVNMC 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*\n2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB pep:*\n3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*\n4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*\n5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*\n6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*\n7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*\n8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*\n9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*\n10: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*\n11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*\n12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*\n13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*\n14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	18	US-09-834-795a-3	Sequence 3, Appl1
2	94	100.0	104	US-09-834-795a-2	Sequence 2, Appl1
3	94	100.0	127	US-09-898-751a-6	Sequence 6, Appl1
4	94	100.0	127	US-09-813-492-2	Sequence 2, Appl1
5	94	100.0	127	US-09-834-795a-1	Sequence 1, Appl1
6	94	100.0	127	US-09-931-381a-2	Sequence 2, Appl1
7	75	79.8	130	US-09-898-751a-10	Sequence 10, Appl1
8	44	46.8	310	US-09-738-626-1198	Sequence 6198, Ap
9	42	44.7	1044	US-09-960-226-3	Sequence 3, Appl1
10	42	44.7	1065	US-09-960-226-2	Sequence 2, Appl1
11	40	42.6	295	US-09-841-132-387	Sequence 387, App
12	40	42.6	302	US-09-757-982-2	Sequence 2, Appl1
13	39.5	42.0	50	US-10-139-876-9	Sequence 9, Appl1
14	39.5	42.0	84	US-10-086-623-13	Sequence 13, Appl1
15	39.5	42.0	146	US-09-921-398-2	Sequence 25, Appl1
16	39.5	42.0	146	US-09-921-398-25	Sequence 25, Appl1
17	39.5	42.0	161	US-09-813-398-7	Sequence 7, Appl1
18	39.5	42.0	205	US-09-921-398-27	Sequence 27, Appl1
19	39.5	42.0	205	US-09-921-398-37	Sequence 37, Appl1

	20	39.5	42.0	241	9	US-09-852-209a-15	Sequence 15, Appl1
	21	39.5	42.0	241	9	US-10-060-523-8	Sequence 8, Appl1
	22	39.5	42.0	241	10	US-09-921-398-29	Sequence 29, Appl1
	23	39.5	42.0	241	10	US-09-749-728b-5	Sequence 5, Appl1
	24	39.5	42.0	241	10	US-09-795-006a-127	Sequence 127, Appl1
	25	39.5	42.0	241	12	US-10-127-551-4	Sequence 4, Appl1
	26	39	41.5	45	10	US-09-925-599-858	Sequence 858, App
	27	39	41.5	437	10	US-09-815-242-5233	Sequence 5233, App
	28	39	41.5	448	10	US-09-815-242-12574	Sequence 12574, A
	29	38	40.4	226	10	US-09-815-242-5900	Sequence 5900, Ap
	30	38	40.4	226	10	US-09-815-242-13156	Sequence 13156, A
	31	38	40.4	240	9	US-09-738-626-3534	Sequence 3534, Ap
	32	38	40.4	424	9	US-10-029-180-104	Sequence 104, App
	33	38	40.4	638	9	US-09-983-204-18	Sequence 18, Appl1
	34	38	40.4	638	9	US-10-133-157-8	Sequence 8, Appl1
	35	38	40.4	1781	10	US-09-738-877-3	Sequence 3, Appl1
	36	37.5	39.9	120	10	US-09-764-864-1211	Sequence 1111, App
	37	37	39.4	90	10	US-09-939-980-497	Sequence 497, App
	38	37	39.4	233	10	US-09-810-808-1	Sequence 1, Appl1
	39	37	39.4	249	9	US-09-764-868-812	Sequence 812, App
	40	37	39.4	271	10	US-09-970-711-50	Sequence 50, Appl1
	41	37	39.4	322	9	US-09-866-570a-14	Sequence 14, Appl1
	42	37	39.4	322	10	US-09-866-572a-14	Sequence 14, Appl1
	43	37	39.4	445	9	US-09-866-570a-52	Sequence 52, Appl1
	44	37	39.4	445	10	US-09-866-572a-52	Sequence 52, Appl1
	45	37	39.4	598	10	US-09-878-764-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-834-795a-3
Sequence 3, Application US/09834795A
Patent No. US20020076710A1

GENERAL INFORMATION:

APPLICANT: Lawrence, Papsidero

APPLICANT: Lyn, Dyster

TITLE OF INVENTION: Detection and Treatment of Breast Cancer

FILE REFERENCE: 3380/11127-US3

CURRENT APPLICATION NUMBER: US/09/834,795A

CURRENT FILING DATE: 2001-04-12

PRIOR APPLICATION NUMBER: 09/146,580

PRIOR FILING DATE: 1998-09-03

PRIOR APPLICATION NUMBER: 60/071,899

PRIOR FILING DATE: 1998-01-20

PRIOR APPLICATION NUMBER: 60/092,155

PRIOR FILING DATE: 1998-07-09

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 18

TYPE: PRT

ORGANISM: Homo sapiens

US-09-834-795a-3

Query Match 100.0% Score 94; DB 10; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e+09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TEVSHHSRRLRLRVNMC 18
DB 1 TEVSHHSRRLRLRVNMC 18

RESULT 2
US-09-834-795a-2

Sequence 2, Application US/09834795A

Patent No. US20020076710A1

GENERAL INFORMATION:

APPLICANT: Lawrence, Papsidero

APPLICANT: Lyn, Dyster

APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 104
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (68)..(68)
OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
US-09-834-795A-2

Query Match 100.0%; Score 94; DB 10; Length 104;
Best Local Similarity 100.0%; Pred. No. 8.8e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TEVSHISRRLRERNVC 18
|||||

Db 9 TEVSHISRRLRERNVC 26

RESULT 3

US-09-898-751A-6
Sequence 6, Application US/09898751A
Patent No. US20020160024A1
GENERAL INFORMATION:
APPLICANT: Oldham, Elizabeth R.
APPLICANT: Soto, Hortensia
APPLICANT: Liu, Ying
APPLICANT: Hudak, Susan A.
APPLICANT: Homey, Bernhard
APPLICANT: Morales, Janine M.
APPLICANT: Kellerman, Sifrid-Almee
APPLICANT: McEvoy, Leslie M.
APPLICANT: Bowman, Edward P.
APPLICANT: Zlotnik, Albert
TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
FILE REFERENCE: DX0882XK
CURRENT APPLICATION NUMBER: US/09/898,751A
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: US09/471,549
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: US60/136,570
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: US60/113,858
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (529)..(529)
OTHER INFORMATION: unknown amino; may be "A", "C", or "G"
US-09-898-751A-6

Query Match 100.0%; Score 94; DB 9; Length 127;

Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TEVSHISRRLRERNVC 18
|||||

Db 32 TEVSHISRRLRERNVC 49

RESULT 4

US-09-813-492-2
Sequence 2, Application US/09813492
Patent No. US20020009735A1
GENERAL INFORMATION:
APPLICANT: Ladow, Mark A.
APPLICANT: Mickanin, Craig Stephen
APPLICANT: Bhatia, Umesh
TITLE OF INVENTION: MAMMARY GLAND CHEMOKINE
FILE REFERENCE: 12345
CURRENT APPLICATION NUMBER: US/09/813,492
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 127
TYPE: PRT
ORGANISM: HUMAN
US-09-813-492-2

Query Match 100.0%; Score 94; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TEVSHISRRLRERNVC 18
|||||

Db 32 TEVSHISRRLRERNVC 49

RESULT 5

US-09-834-795A-1
Sequence 1, Application US/09834795A
Patent No. US20020076710A1
GENERAL INFORMATION:
APPLICANT: Lawrence, Papsidero
APPLICANT: Lyn, Dyster
APPLICANT: Jana, Frustaci
TITLE OF INVENTION: Detection and Treatment of Breast Cancer
FILE REFERENCE: 3380/11127-US3
CURRENT APPLICATION NUMBER: US/09/834,795A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 09/146,580
PRIOR FILING DATE: 1998-09-03
PRIOR APPLICATION NUMBER: 60/071,899
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 60/092,155
PRIOR FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 35
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
NAME/KEY: UNSURE
LOCATION: (91)..(91)
OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
US-09-834-795A-1

Query Match 100.0%; Score 94; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TEVSHHISRRLERYNC 18
Db 32 TEVSHHISRRLERYNC 49

RESULT 6
US-09-931-381A-2

; Sequence 2, Application US/09931381A
; Patent No. US20020137107A1
; GENERAL INFORMATION:
; APPLICANT: Butcher, Eugene C.
; APPLICANT: Kunkel, Eric J.
; APPLICANT: Pan, Junliang
; APPLICANT: Soler-Ferran, Dulce
; TITLE OF INVENTION: Method for Identifying Agents Which
; TITLE OF INVENTION: Modulate Chemokine "Mec"-Induced Functions of CCR3 and/or,
; TITLE OF INVENTION: CCR3
; FILE REFERENCE: 1855.2010-003
; CURRENT APPLICATION NUMBER: US/09/931,381A
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: U.S. 09/638,914
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(22)
US-09-931-381A-2

Query Match 100.0%; Score 94; DB 10; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TEVSHHISRRLERYNC 18
Db 32 TEVSHHISRRLERYNC 49

RESULT 7
US-09-898-751A-10

; Sequence 10, Application US/09898751A
; Patent No. US20020160024A1
; GENERAL INFORMATION:
; APPLICANT: Oldham, Elizabeth R.
; APPLICANT: Solo, Hortensia
; APPLICANT: Liu, Ying
; APPLICANT: Hudak, Susan A.
; APPLICANT: Morales, Janine M.
; APPLICANT: Kellerman, Strid-Almege
; APPLICANT: McEvoy, Leslie M.
; APPLICANT: Bowman, Edward P.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: CHEMOKINE AND RECEPTOR USES; COMPOSITIONS; METHODS
; FILE REFERENCE: DX0882X
; CURRENT APPLICATION NUMBER: US/09/898,751A
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: US09/471,549
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US60/136,570
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: US60/113,858
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 130
; TYPE: PRT

; ORGANISM: Mus musculus
US-09-898-751A-10

Query Match 79.8%; Score 75; DB 9; Length 130;
Best Local Similarity 77.8%; Pred. No. 1.5e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 TEVSHHISRRLERYNC 18
Db 32 TEVSHHISRRLERYNC 49

RESULT 8
US-09-738-626-6198

; Sequence 6198, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6198
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6198

Query Match 46.8%; Score 44; DB 9; Length 310;
Best Local Similarity 44.4%; Pred. No. 4.8;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 TEVSHHISRRLERYNC 18
Db 144 TEVSHHISRRLERYNC 161

RESULT 9
US-09-960-226-3

; Sequence 3, Application US/09960226
; Patent No. US20020172670A1
; GENERAL INFORMATION:
; APPLICANT: Kuntz, David
; APPLICANT: Rose, Douglas
; APPLICANT: Van Den Elsen, Jean
; TITLE OF INVENTION: MANNOSIDASE STRUCTURES
; FILE REFERENCE: 12243.19US01
; CURRENT APPLICATION NUMBER: US/09/960,226
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/263,458
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: US 60/234,879
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

LENGTH: 1044
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-960-226-3

Query Match 44.7%; Score 42; DB 9; Length 1044;
Best Local Similarity 61.5%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HISRRLLERVNC 18
|||:|||||
Db 889 HIRLVLEKVNMC 901

RESULT 10
US-09-960-226-2
Sequence 2, Application US/09960226
Patent No. US20020172670A1
GENERAL INFORMATION:
APPLICANT: Rose, David
APPLICANT: Kuntz, Douglas
APPLICANT: Van Den Elsen, Jean
TITLE OF INVENTION: MANNOSIDASE STRUCTURES
FILE REFERENCE: 12243.19USU1
CURRENT APPLICATION NUMBER: US/09/960,226
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/263,458
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: US 60/234,879
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1065
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
NAME/KEY: misc.feature
LOCATION: (542)..(542)
OTHER INFORMATION: The 'Xaa' at location 542 stands for Pro, or Ser.
US-09-960-226-2

Query Match 44.7%; Score 42; DB 9; Length 1065;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 HISRRLLERVNC 18
|||:|||||
Db 910 HIRLVLEKVNMC 922

RESULT 11
US-09-841-132-387
Sequence 387, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhalla, Ajay
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 387
LENGTH: 295
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-841-132-387
Query Match 42.6%; Score 40; DB 10; Length 295;

Best Local Similarity 40.0%; Pred. No. 20;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TEVSHISRLLERV 15
||||:|||||
Db 14 THQSHLDKRYERL 28

RESULT 12
US-09-757-982-2
Sequence 2, Application US/09757982
Patent No. US20020094559A1
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/757,982
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 09/163,115
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-757-982-2

Query Match 42.6%; Score 40; DB 10; Length 302;
Best Local Similarity 53.3%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHHSRLLERVNC 18
|||:|||||
Db 260 SDHYSEELRQLVNMC 274

RESULT 13
US-10-139-876-9
Sequence 9, Application US/10139876
Patent No. US20020123481A1
GENERAL INFORMATION:
APPLICANT: Oliviero, Salvatore
TITLE OF INVENTION: C-Fos Induced Growth Factor (Figf) And Dna Encoding Same
FILE REFERENCE: 35784/205112
CURRENT APPLICATION NUMBER: US/10/139,876
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 09/043,476
PRIOR FILING DATE: 1998-03-18
PRIOR APPLICATION NUMBER: PCT/IB96/0113
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: GB9612368.2
PRIOR FILING DATE: 1996-06-13
PRIOR APPLICATION NUMBER: GB9519928.7
PRIOR FILING DATE: 1995-09-29
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 50
TYPE: PRT
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: mammalian
NAME/KEY: PEPTIDE
LOCATION: (1)..(50)
OTHER INFORMATION: segment of PDGF-B
US-10-139-876-9

Query Match 42.0%; Score 39.5; DB 12; Length 50;
Best Local Similarity 62.5%; Pred. No. 3.4;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 TEVSHHSRRLERVN 16
||| |||||:|
DB 18 TEV-FEISRLIDRTN 32

RESULT 14

US-10-086-623-13
Sequence 13, Application US/10086623

Patent No. US20020164710A1

GENERAL INFORMATION:

APPLICANT: ERIKSSON, Ulf

APPLICANT: AASE, Karin

APPLICANT: LI, Xuri

APPLICANT: PONTEN, Annica

APPLICANT: UTELA, Marko

APPLICANT: ALITALO, Kari

APPLICANT: OESTMAN, Arne

APPLICANT: HELDIN, Carl-Henrik

TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR D, DNA CODING THEREFOR AND USES TH

FILE REFERENCE: 1064/44833C2

CURRENT FILING DATE: 2000-03-04

PRIOR APPLICATION NUMBER: US 60/107,852

PRIOR FILING DATE: 1998-11-10

PRIOR APPLICATION NUMBER: US 60/113,997

PRIOR FILING DATE: 1998-12-28

PRIOR APPLICATION NUMBER: US 60/150,604

PRIOR FILING DATE: 1999-08-26

PRIOR APPLICATION NUMBER: US 60/157,108

PRIOR FILING DATE: 1999-10-04

PRIOR APPLICATION NUMBER: US 60/157,756

PRIOR FILING DATE: 1999-10-05

PRIOR APPLICATION NUMBER: US 09/438,046

PRIOR FILING DATE: 1999-11-10

PRIOR APPLICATION NUMBER: US 09/691,200

PRIOR FILING DATE: 2000-10-19

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn version 3.1

SEQ ID NO 13

LENGTH: 84

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: PDGF/VEGF-homology domain of PDGF-B

US-10-086-623-13

Query Match 42.0%; Score 39.5; DB 9; Length 84;

Best Local Similarity 62.5%; Pred. No. 6.1;

Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

OY 1 TEVSHHSRRLERVN 16
||| |||||:|

DB 5 TEV-FEISRLIDRTN 19

RESULT 15

US-09-921-398-2

Sequence 2, Application US/09921398

Patent No. US20020055169A1

GENERAL INFORMATION:

APPLICANT: Tekamp-Olson, Patricia

TITLE OF INVENTION: METHOD FOR EXPRESSION OF HETEROLOGOUS

PROTEINS IN YEAST

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell Seltzer IP Group of Alston & Bird, LLP

STREET: 3605 Glenwood Ave, Suite 310

CITY: Raleigh

STATE: NC

COUNTRY: US

ZIP: 27622

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/921,398
FILING DATE: 02-Aug-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: 5784-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919 420 2202
TELEFAX: 919 881 3175
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 146 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-921-398-2

Query Match 42.0%; Score 39.5; DB 10; Length 146;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
OY 1 TEVSHHSRRLERVN 16
||| |||||:|
DB 57 TEV-FEISRLIDRTN 71

Search completed: January 14, 2003, 18:27:28
Job time : 3.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:40:55 ; Search time 11.5 Seconds

(without alignments)
208.566 Million cell updates/sec

Title: US-09-834-794A-3

Perfect score: 94

Sequence: 1 TEVSHHSRRLRLRRVNC 18

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	18	20	AAV29094
2	94	100.0	97	22	AA850353
3	94	100.0	104	20	AAV29093
4	94	100.0	108	20	AAV11558
5	94	100.0	127	19	AAW60649
6	94	100.0	127	20	AAV29092
7	94	100.0	127	21	AA801450
8	94	100.0	127	21	AA803001
9	94	100.0	127	22	AA859175
10	94	100.0	127	23	AB807728

11	94	100.0	127	23	AB808268
12	94	100.0	147	22	AAW25479
13	88	93.6	106	23	AA847869
14	88	93.6	127	23	AA847868
15	75	79.8	130	21	AA801452
16	45	47.9	681	22	AB828711
17	44	46.8	310	22	AA832444
18	43	45.7	581	23	AB852301
19	42	44.7	388	22	AB855417
20	42	44.7	1032	22	AB809187
21	42	44.7	1108	22	AB863713
22	41	43.6	104	13	AA825441
23	41	43.6	104	17	AA87513
24	41	43.6	182	22	AB862301
25	41	43.6	264	22	AA868590
26	41	43.6	470	23	ABP10175
27	41	43.6	476	23	ABP26844
28	41	43.6	1069	22	AB855308
29	41	43.6	1381	22	AAU68547
30	40	42.6	108	22	AA824928
31	40	42.6	132	21	AA824928
32	40	42.6	132	21	AAV79156
33	40	42.6	158	21	AA824927
34	40	42.6	246	20	AAV3575
35	40	42.6	293	20	AAV13739
36	40	42.6	293	21	AAV80299
37	40	42.6	295	23	AB84259
38	40	42.6	300	20	AAV35008
39	40	42.6	302	21	AAV84320
40	40	42.6	302	21	AAV59143
41	40	42.6	302	22	AAV71956
42	40	42.6	311	22	AAV5475
43	40	42.6	429	20	AAW3213
44	40	42.6	459	21	AAV74413
45	40	42.6	1428	22	AB868557

ALIGNMENTS

RESULT 1
ID AAV29094
AAV29094 standard; peptide: 18 AA.
AC AAV29094;
DT 29-SEP-1999 (first entry)
DE Human mammary associated chemokine (MACK) protein fragment MACK A.
KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
KW Inflammation; infection; mastitis; benign cystitis; hyperplasia;
KW mammary associated chemokine; MACK; epitope.
OS Homo sapiens.
XX MO9936540-A1.
XX 22-JUL-1999.
XX 12-JAN-1999; 99WO-US00651.
XX 09-JUL-1998; 98US-0092155.
XX 20-JAN-1998; 98US-0071899.
XX (CODON-) CODON DIAGNOSTICS LLC.
XX Dyster LM, Frustaci JM, Papsidero LD;
XX WPI: 1999-458469/38.
XX A mammary associated chemokine and related polynucleotides, useful
XX for detection and treatment of breast disease, especially cancer

XX Claim 8; Page 66; 76pp; English.

PS The invention provides an isolated human chemokine, which is

CC preferentially expressed in breast tissue or detected in breast milk. An

CC antibody that recognizes the novel chemokine, or a chemokine-derived

CC antigenic peptide, can be used to treat breast disease in a patient. A

CC peptide, which binds to a cellular receptor for the chemokine, can also

CC be used to treat breast disease. Antigenic peptides of the chemokine

CC polynucleotide sequences and the chemokine protein can be detected in

CC samples with primers, probes and antibodies using standard techniques.

CC This is useful for detecting breast disease. Other breast diseases that

CC may be treated or detected with the chemokine and its encoding

CC polynucleotides include inflammations, infections, mastitis, benign

CC cystitis, and benign hyperplasias as well as other malignancies.

CC Sequences AAY29094-96 represent fragments of the human mammary

CC associated chemokine (MACK) protein. These fragments act as antigenic

CC peptide epitopes against which rabbit antisera can be raised.

XX

SQ Sequence 18 AA;

Query Match 100.0%; Score 94; DB 20; Length 18;

Best Local Similarity 100.0%; Pred. No. 5.6e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TEVSHHSRRLRLRVNMC 18

Db 1 TEVSHHSRRLRLRVNMC 18

RESULT 2

AAB50353

ID AAB50353 standard; Protein; 97 AA.

XX

AC AAB50353;

XX

DT 09-MAR-2001 (first entry)

XX

DE Human RIntIntIn protein.

XX

KW Human; RIntIntIn; antiallergic; antiasthmatic; antiarteriosclerotic;

KW immunosuppressive; cytostatic; osteopathic; antimicrobial; antiporiatic;

KW antiallergic; vasotropic; vaccine; chemokine; allergy; asthma;

KW atherosclerosis; autoimmune disease; cancer; osteoarthritis; infection;

KW psoriasis; rheumatoid arthritis; vascular disease.

OS

XX Homo sapiens.

XX

PN WO200073447-A1.

XX

PD 07-DEC-2000.

XX

PF 24-MAY-2000; 2000MO-US14197.

XX

PR 28-MAY-1999; 9905-0321651.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.

XX

PI Appelbaum E, Vawter L;

XX

DR WPI: 2001-061533/07.

XX

DR N-PSDB; AAC89722.

XX

PT RIntIntIn polypeptides and polynucleotides useful as a diagnostic

PT tool and for treating and preventing various diseases e.g. autoimmune

PT diseases, cancer, infections, bone disorders etc

XX

PS Claim 1; Page 27; 30pp; English.

XX

CC The present sequence is a chemokine designated RIntIntIn. The

CC RIntIntIn polynucleotide is useful as a diagnostic reagent through

CC detecting mutations in the associated gene. The polynucleotide and

CC polypeptide are useful for screening agonists and antagonists of

CC RIntIntIn and are also useful as vaccines in treating various

CC diseases such as allergies, asthma, atherosclerosis, autoimmune

CC disease, cancer, osteoarthritis, various infections, psoriasis,

CC rheumatoid arthritis, and vascular disease.

XX

SQ Sequence 97 AA;

Query Match 100.0%; Score 94; DB 22; Length 97;

Best Local Similarity 100.0%; Pred. No. 3.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TEVSHHSRRLRLRVNMC 18

Db 32 TEVSHHSRRLRLRVNMC 49

RESULT 3

AAY29093

ID AAY29093 standard; Protein; 104 AA.

XX

AC AAY29093;

XX

DT 29-SEP-1999 (first entry)

XX

DE Human mammary associated chemokine (MACK) mature protein sequence.

XX

KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;

KW inflammation; infection; mastitis; benign cystitis; hyperplasia;

KW mammary associated chemokine; MACK.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT MISC-difference 47

FT /label= unknown

FT /note= "encoded by NGA"

FT MISC-difference 68

FT /label= unknown

FT /note= "encoded by AAN"

XX

PN WO936540-A1.

XX

PD 22-JUL-1999.

XX

PF 12-JAN-1999; 99WO-US00651.

XX

PR 09-JUL-1998; 98US-0092155.

XX

PR 20-JAN-1998; 98US-0071899.

XX

PA (CODO-) CODON DIAGNOSTICS LLC.

XX

PI Dyster LM, Frustaci JM, Papsidero LD;

XX

DR WPI: 1999-458469/38.

XX

DR N-PSDB; AAX89389, AAX89390.

XX

PT A mammary associated chemokine and related polynucleotides, useful

PT for detection and treatment of breast disease, especially cancer

XX

PS Claim 5; Page 66; 76pp; English.

XX

CC The invention provides an isolated human chemokine, which is

CC preferentially expressed in breast tissue or detected in breast milk. An

CC antibody that recognizes the novel chemokine, or a chemokine-derived

CC antigenic peptide, can be used to treat breast disease in a patient. A

CC peptide, which binds to a cellular receptor for the chemokine, can also

CC be used to treat breast disease. Antigenic peptides of the chemokine can

CC be used to vaccinate patients against breast disease. The chemokine

CC polynucleotide sequences and the chemokine protein can be detected in

CC samples with primers, probes and antibodies using standard techniques.

CC This is useful for detecting breast disease. Other breast diseases that

CC may be treated or detected with the chemokine and its encoding

CC polynucleotides include inflammations, infections, mastitis, benign
 CC cystitis, and benign hyperplasias as well as other malignancies. The
 CC present sequence represents the human mammary associated chemokine (MACK)
 CC mature protein sequence.

XX Sequence 104 AA;

Query Match 100.0%; Score 94; DB 20; Length 104;
 Best Local Similarity 100.0%; Pred. No. 3.7e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TEVSHHSRRLRRVNC 18
 DB 9 TEVSHHSRRLRRVNC 26

RESULT 4

AAV11658
 ID AAV11658 standard; Protein; 108 AA.

AC AAV11658;

DT 16-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO:310.

XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 XX forensic; gene therapy; chromosome mapping; signal peptide;
 XX upstream regulatory sequence; cytokine activity; cell proliferation;
 XX differentiation; hematopoiesis regulation; tissue growth regulation;
 XX reproductive hormone regulation; chemotactic; haemostatic;
 XX thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

XX MO9906439-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; 98MO-IB01233.

PR 01-AUG-1997; 97US-0904468.

XX (GEST) GENSET.

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

XX WPI; 1999-153700/13.

DR N-PSDB; AAX40376.

XX New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from liver, lung, large intestine, colon,
 PT thyroid and pancreas tissue

PS Claim 27; Page 384-385; 398pp; English.

XX AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAV11533 to
 CC AAV11679, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, hematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

XX Sequence 108 AA;

Query Match 100.0%; Score 94; DB 20; Length 108;
 Best Local Similarity 100.0%; Pred. No. 3.8e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TEVSHHSRRLRRVNC 18
 DB 32 TEVSHHSRRLRRVNC 49

RESULT 5

AAW60649
 ID AAW60649 standard; Protein; 127 AA.

AC AAW60649;

DT 12-OCT-1998 (first entry)

DE Human DVIC-1 C-C chemokine.

XX DVIC-1; DNAX Vlc-1; C-C chemokine; cytokine; human; immune system;
 XX cancer; cell proliferation; therapy; diagnosis.

OS Homo sapiens.

XX Key location/Qualifiers

FT Peptide 1..22 /label= Sig_peptide

FT Protein 23..127 /label= Mat_protein

FT /note= "Claim 1"

XX MO9823750-A2.

PD 04-JUN-1998.

PF 26-NOV-1997; 97MO-US21092.

PR 05-DEC-1996; 96US-0761071.

XX 27-NOV-1996; 96US-0031805.

XX (SCHE) SCHERING CORP.

PI Hedrick JA, Morales J, Vicari A, Zlotnik A;

XX WPI; 1998-322730/28.

DR N-PSDB; AAV38291.

XX DVIC-1 and DGMCC chemokines - useful for developing products for
 PT treating abnormal physiology or development, e.g. cancerous or
 PT degenerative conditions

PS Claim 1; Page 59-60; 71pp; English.

XX This polypeptide comprises human DNAX Vlc-1 (DVIC-1), a novel C-C
 CC chemokine, the mature portion of which is claimed. The amino acid
 CC sequence was deduced from a cDNA clone (see AAV38291). An alternative
 CC longer transcript (see AAW60652) for human DVIC-1 is also disclosed.
 CC Also claimed is novel human DNAX Groin Wound expressed CC chemokine
 CC (DGMCC) (see AAW60649) mature protein, as well as expression vectors
 CC and host cells. DVIC-1 and DGMCC play a role in the regulation or
 CC development of neuronal or haematopoietic cells, e.g. lymphoid
 CC cells, which affect immunological responses. They can be used in
 CC the treatment of conditions associated with abnormal physiology or
 CC development, including abnormal proliferation, e.g. cancerous
 CC conditions or degenerative conditions. Abnormal proliferation,
 CC regeneration, degeneration, and atrophy may be modulated by
 CC appropriate therapeutic treatment using products of the invention.
 CC The products can also be used for detection, diagnosis and drug
 CC screening.

SO Sequence 127 AA:

Query Match 100.0%; Score 94; DB 19; Length 127;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLRLRVNMC 18
 |||
 32 TEVSHHSRRLRLRVNMC 49

Db

RESULT 6
 AAY29092
 ID AAY29092 standard; Protein; 127 AA.

AC AAY29092;
 XX
 DT 29-SEP-1999 (first entry)

DE Human mammary associated chemokine (MACK) protein.
 XX
 KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
 KM Inflammation; infection; mastitis; benign cystitis; hyperplasia;
 KM mammary associated chemokine; MACK.
 XX
 OS Homo sapiens.

XX
 FH Key
 FT Peptide 1..23
 FT /note="signal peptide"
 FT Protein 24..127
 FT /note="mature protein (AAY29093)"
 FT Region 32..49
 FT /note="antigenic peptide epitope MACK A (AAY29094)"
 FT Misc-difference 70
 FT /label="unknown
 FT /note="encoded by NGA"
 FT Misc-difference 91
 FT /label="unknown
 FT /note="encoded by AAN"
 FT Region 92..107
 FT /note="antigenic peptide epitope MACK B (AAY29095)"
 FT Region 109..127
 FT /note="antigenic peptide epitope MACK C (AAY29096)"

XX
 PN WO936540-A1.
 XX
 PD 22-JUL-1999.
 XX
 PF 12-JAN-1999; 99MO-US00651.
 XX
 PR 09-JUL-1998; 98US-0092155.
 PR 20-JAN-1998; 98US-0071899.
 XX
 PA (CODO-) CODON DIAGNOSTICS LLC.
 XX
 PI Dyster LM, Frustaci JM, Papsidero LD;
 PI
 DR MPI; 1999-458469/38.
 DR N-PSDB; AAX89389, AAX89390.
 XX
 PT A mammary associated chemokine and related polynucleotides, useful
 PT for detection and treatment of breast disease, especially cancer
 XX
 PS Claim 4; Page 48; 76pp; English.
 XX
 CC The invention provides an isolated human chemokine, which is
 CC preferentially expressed in breast tissue or detected in breast milk. An
 CC antibody that recognizes the novel chemokine, or a chemokine-derived
 CC antigenic peptide, can be used to treat breast disease in a patient. A
 CC peptide, which binds to a cellular receptor for the chemokine, can also
 CC be used to treat breast disease. Antigenic peptides of the chemokine can
 CC be used to vaccinate patients against breast disease. The chemokine

CC polynucleotide sequences and the chemokine protein can be detected in
 CC samples with primers, probes and antibodies using standard techniques.
 CC This is useful for detecting breast disease. Other breast diseases that
 CC may be treated or detected with the chemokine and its encoding
 CC polynucleotides include inflammations, infections, mastitis, benign
 CC cystitis, and benign hyperplasias as well as other malignancies. The
 CC present sequence represents the amino acid sequence of the human mammary
 CC associated chemokine (MACK) protein.

SO Sequence 127 AA:

Query Match 100.0%; Score 94; DB 20; Length 127;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLRLRVNMC 18
 |||
 32 TEVSHHSRRLRLRVNMC 49

Db

RESULT 7
 AAB01450
 ID AAB01450 standard; Protein; 127 AA.

AC AAB01450;
 XX
 DT 20-OCT-2000 (first entry)

DE Primate VIC protein sequence.
 XX
 KW Cutaneous T-cell attracting chemokine; CTACK; skin; cell movement;
 KW migration; vasoreactive intestinal contractor; VIC; GPR2; agonist;
 KW antagonist; antibody; immunological condition; mutain.
 XX
 OS Homo sapiens.
 XX
 PN WO200038713-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 23-DEC-1999; 99MO-US30819.
 XX
 PR 24-DEC-1998; 98US-0113858.
 PR 27-MAY-1999; 99US-0322580.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Wang W, Oldham ER, Soto H, Lui Y, Hudak SA, Homey B, Morales JM;
 PI Kellermann S, McEvoy LM, Zlotnik A;
 PI
 DR MPI; 2000-465633/40.
 DR N-PSDB; AAA47545.
 XX
 PT Modulating cell movement within the skin, useful for treating
 PT immunological skin conditions or diseases comprising administering T
 PT cell-attracting chemokine or vasoreactive intestinal contractor chemokine
 PT agonists or antagonists
 XX
 PS Example 3; Page 69; 79pp; English.
 XX
 CC Modulating movement of a cell within or to the skin of a mammal can
 CC be achieved by administering an antagonist or agonist of cutaneous T
 CC cell-attracting chemokine (CTACK) or vasoreactive intestinal contractor
 CC (VIC) chemokine. The antagonist is selected from a mixture of natural
 CC CTACK or VIC, an antibody which neutralises CTACK or VIC or an
 CC antibody which block GPR2 ligand binding. The CTACK or VIC agonists
 CC or antagonists are useful for treating medical conditions or diseases
 CC associated with immunological conditions of the skin.

SO Sequence 127 AA:

Query Match 100.0%; Score 94; DB 21; Length 127;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;

Matches 18: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 1 TEVSHHSRLLERYNMC 18
 |||||
 Db 32 TEVSHHSRLLERYNMC 49

RESULT 8
 AAB03001
 ID AAB03001 standard; Protein: 127 AA.

AC AAB03001;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Human growth factor related molecule GFRP-2.
 XX
 KW Human GFRP-2; growth factor related molecule; breast tissue;
 KW breast tumour; CC chemokine; hTBECK homologue; developmental disorder;
 KW cell proliferative disorder; immune disorder; reproductive disorder;
 KW cardiovascular disorder; bacterial infection; viral; fungal; parasitic;
 KW cancer; allergy; asthma; arteriosclerosis; therapy; diagnosis.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..19
 FT /note= "Putative signal peptide"
 FT Peptide 1..22
 FT /note= "Putative signal peptide"
 FT Region 30..74
 FT /note= "This region has strong similarity to the CC
 chemokine consensus sequence"
 FT Modified-site 39
 FT /note= "Phosphorylated by protein kinase C"
 FT Modified-site 78
 FT /note= "N-glycosylated"
 FT Modified-site 80
 FT /note= "Phosphorylated by protein kinase C"
 FT Modified-site 110
 FT /note= "Phosphorylated by CAMP/CGMP-dependent protein
 kinase or protein kinase C"
 FT
 FT
 PN WO200024774-A2.
 XX
 PD 04-MAY-2000.
 XX
 PE 28-OCT-1999; 99WO-US25458.
 XX
 PR 28-OCT-1998; 98US-0181711.
 PR 11-DEC-1998; 98US-0209547.
 PR 17-MAY-1999; 99US-0313457.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Tang YT, Yue H, Hillman JL, Corley NC, Guegler KJ, Baughn MR;
 PI Au-Young J;
 XX
 DR WPI: 2000-350695/30.
 DR N-PSDB: AAA52456.
 XX
 XX Human growth factor related molecule protein useful for the diagnosis
 PT and treatment of disorders associated with its activity including
 PT developmental, cell proliferative, immune, reproductive and
 PT cardiovascular disorders and infections -
 XX
 PS Claim 1; Fig 2; 80pp; English.
 XX
 CC This sequence represents human growth factor related molecule GFRP-2.
 CC cDNA encoding GFRP-1 was initially identified in a breast tissue
 CC cDNA library, and the present sequence is encoded by a consensus
 CC cDNA derived from several overlapping and/or extended cDNA clones.
 CC GFRP-2 is probably a CC chemokine and has chemical and structural

CC homology with hTBECK (20% identity). GFRP-2 and hTBECK are both basic
 CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
 CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
 CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),
 CC nucleotides encoding them (AA52455-AA52458), GFRP agonists and
 CC antagonists may be used to treat a wide variety of diseases associated
 CC with increased or decreased expression or activity of GFRP proteins.
 CC Conditions which may be treated include developmental disorders, cell
 CC proliferative disorders (e.g., cancers), immune disorders (e.g.,
 CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
 CC disorders), cardiovascular disorders (e.g., arteriosclerosis) and
 CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
 CC proteins and nucleotides can be used in the diagnosis of such disorders.
 XX
 XX Sequence 127 AA;
 XX
 XX
 XX Query Match 100.0%; Score 94; DB 21; Length 127;
 XX Best Local Similarity 100.0%; Pred. No. 4; 3e-08;
 XX Matches 18: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 1 TEVSHHSRLLERYNMC 18
 |||||
 Db 32 TEVSHHSRLLERYNMC 49

RESULT 9
 AAB69175
 ID AAB69175 standard; Protein: 127 AA.

AC AAB69175;
 XX
 DT 26-APR-2001 (first entry)
 XX
 DE Human G-protein coupled receptor GPR27 protein SEQ ID NO:2.
 XX
 KW Human; G-protein coupled receptor; GPR27; antimicrobial; analgesic;
 KW cytosolic; antidiabetic; anorectic; antiaesthmic; antiparkinsonian;
 KW cardiant; hypertensive; hypotensive; diuretic; osteopathic; antitumor;
 KW cerebroprotective; antiallergic; antileptic; tranquilliser; nootropic;
 KW antidepressant; neuroleptic; anticonvulsant; vaccine; gene therapy;
 KW infection; pain; cancer; diabetes; obesity; anorexia; bulimia; asthma;
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; vomiting;
 KW psychotic disorder; neurological disorder; dyskinesia.
 XX
 OS Homo sapiens.
 XX
 PN WO200107482-A1.
 XX
 PD 01-FEB-2001.
 XX
 PE 21-JUL-2000; 2000WO-US19855.
 XX
 PR 27-JUL-1999; 99US-0361564.
 XX
 PA (SMIRK) SMITHKLINE BEECHAM CORP.
 PA (SMIRK) SMITHKLINE BEECHAM PLC.
 XX
 PI Teßta TT, Vawter L;
 XX
 DR WPI: 2001-191359/19.
 DR N-PSDB: AAF59231.
 XX
 XX New G-protein coupled receptor GPR27 polypeptides and polynucleotides,
 PT useful as vaccines or for treating diseases, e.g. infections, pain,
 PT cancers, urinary retention, osteoporosis, stroke, psychotic and
 PT neurological disorders -
 XX
 PS Claim 1; Page 29-30; 31pp; English.
 XX
 CC The present sequence represents a human G-protein coupled receptor,
 CC designated GPR27. GPR27 has antimicrobial, analgesic, cytosolic,

CC invention relates to a novel method for regulating a tumour or
CC adverse bodily reaction, comprising providing a therapeutic composition
CC having a mammary gland chemokine polypeptide. The polypeptide of the
CC invention has cytostatic and antiinflammatory activity. The method of the
CC invention is useful for regulating a tumour or adverse bodily reaction.
CC The invention also provides a method useful for detecting a tumour using
CC a probe comprising the polynucleotide or an antibody to the MEC. The
CC adverse bodily reactions include cancer and inflammation.
XX
SQ Sequence 127 AA;
Query Match 100.0%; Score 94; DB 23; Length 127;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DY 1 TEVSHHSRRLRLRYNMC 18
|||
DB 32 TEVSHHSRRLRLRYNMC 49
RESULT 12
AAM25479
ID AAM25479 standard; Protein: 147 AA.
AC AAM25479;
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:994.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
KW anti-HIV; fungicide; antitumagen; cardiovascular; antianaemic; anemia;
KW antileptogenic; haemostatic; vulnery; antilept; osteopathic; eczema;
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antiapophytic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
OS Homo sapiens.
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000MO-US35017.
XX
PR 23-DEC-1999; 9905-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR N-PSDB; AAH99420.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
PS Claim 20; Page 207; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25663. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;

CC antirheumatic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; vitruide; anti-HIV; fungicide; antitumagen;
CC cardiovascular; antianaemic; antileptogenic; haemostatic; vulnery;
CC antilept; osteopathic; dermatological; antiallergic; antiaesthetic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmune, genetic diseases, haematopoietic disorders,
CC anemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
SQ Sequence 147 AA;
Query Match 100.0%; Score 94; DB 23; Length 147;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DY 1 TEVSHHSRRLRLRYNMC 18
|||
DB 52 TEVSHHSRRLRLRYNMC 69
RESULT 13
AAB47869
ID AAB47869 standard; Protein: 106 AA.
AC AAB47869;
DT 15-APR-2002 (first entry)
XX
DE Mature placenta-derived C-C chemokine.
XX
KW Chemokine; placenta-derived; C-C chemokine; PLACC; inflammation;
KW immune response; cell proliferation; asthma; malignancy; infection;
KW adult respiratory distress syndrome; rheumatoid arthritis;
KW lupus erythematosus; psoriasis; osteoarthritis; glomerulonephritis;
KW osteoporosis; dermatomyositis; polymyositis; Addison's Disease;
KW Graves Disease; Crohn's Disease; irritable bowel syndrome;
KW atrophic gastritis; graft versus host disease; myasthenia gravis;
KW multiple sclerosis; autoimmune thyroiditis; ulcerative colitis;
KW atherosclerosis; pancreatitis.
XX
OS Homo sapiens.
XX
PN WO200192301-A2.
XX
PD 06-DEC-2001.
XX
PF 23-MAY-2001; 2001MO-US16599.
XX
PR 26-MAY-2000; 2000US-207578P.
XX
PA (INDV) UNIV INDIANA ADVANCED RES & TECHNOLOGY.
XX
PI Hromas RA;
XX
DR N-PSDB; AAI72313.
XX
PT New placenta-derived human C-C chemokine, useful for modulating an
XX inflammatory or immune response, particularly for treating asthma,
XX psoriasis, osteoarthritis, Graves Disease, Crohn's Disease or graft
XX versus host disease -
XX
PS Disclosure; Fig 1; 97pp; English.

XX This sequence shows a chemokine protein, which is a placenta-derived
CC C-C chemokine (PLACC). The PLACC polypeptide or the chemokine-encoding
CC polynucleotide, are useful for modulating an inflammatory/immune
CC response, modulating the activity of PLACC and screening for modulators,
CC and in stimulating cell proliferation. Particularly these are useful
CC for treating asthma, adult respiratory distress syndrome, rheumatoid
CC arthritis, lupus erythematosus, psoriasis, osteoarthritis,
CC glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
CC Addison's disease, Graves Disease, Crohn's Disease, irritable bowel
CC syndrome, atrophic gastritis, graft versus host disease, myasthenia
CC gravis, multiple sclerosis, autoimmune thyroiditis, ulcerative colitis,
CC atherosclerosis or pancreatitis. Increasing the PLACC activity in the
CC subject is useful where the subject may be exhibiting signs of a
CC malignancy or infection or have an immunodeficiency.
XX

SQ Sequence 106 AA;

Query Match 93.6%; Score 88; DB 23; Length 106;
Best Local Similarity 94.4%; Pred. No. 3.8e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TEVSHHSRRRLRVNMC 18
|||||||
DB 11 TEVSHHSRRRLRVNMC 28

RESULT 14
AAB47868
ID AAB47868 standard; Protein; 127 AA.
XX
AC AAB47868;
XX
DT 15-APR-2002 (first entry)
XX
DE Full length placenta-derived C-C chemokine.
XX
KW Chemokine; placenta-derived; C-C chemokine; PLACC; inflammation;
KW Immune response; cell proliferation; asthma; malignancy; infection;
KW adult respiratory distress syndrome; rheumatoid arthritis;
KW lupus erythematosus; psoriasis; osteoarthritis; glomerulonephritis;
KW osteoporosis; dermatomyositis; polymyositis; Addison's Disease;
KW Graves Disease; Crohn's Disease; irritable bowel syndrome;
KW atrophic gastritis; graft versus host disease; myasthenia gravis;
KW multiple sclerosis; autoimmune thyroiditis; ulcerative colitis;
KW atherosclerosis; pancreatitis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT /Label= signal_peptide
FT Protein 22..127
FT /Label= C-C_chemokine
XX
PN WO200192301-A2.
XX
PD 06-DEC-2001.
XX
PF 23-MAY-2001; 2001WO-US16599.
XX
PR 26-MAY-2000; 2000US-207578P.
XX
PA (INDV) UNITV INDIANA ADVANCED RES & TECHNOLOGY.
XX
PI Hromas RA;
XX
PI WPI: 2002-154522/20.
XX
DR N-PSDB: AAI72313.
XX
PT New placenta-derived human C-C chemokine, useful for modulating an
PT inflammatory or immune response, particularly for treating asthma,
PT psoriasis, osteoarthritis, Graves Disease, Crohn's Disease or graft

PT versus host disease -
XX
PS Claim 1; Page 90-91; 97pp; English.
XX

CC This sequence shows a chemokine protein, which is a placenta-derived
CC C-C chemokine (PLACC). The PLACC polypeptide or the chemokine-encoding
CC polynucleotide, are useful for modulating an inflammatory/immune
CC response, modulating the activity of PLACC and screening for modulators,
CC and in stimulating cell proliferation. Particularly these are useful
CC for treating asthma, adult respiratory distress syndrome, rheumatoid
CC arthritis, lupus erythematosus, psoriasis, osteoarthritis,
CC glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
CC Addison's disease, Graves Disease, Crohn's Disease, irritable bowel
CC syndrome, atrophic gastritis, graft versus host disease, myasthenia
CC gravis, multiple sclerosis, autoimmune thyroiditis, ulcerative colitis,
CC atherosclerosis or pancreatitis. Increasing the PLACC activity in the
CC subject is useful where the subject may be exhibiting signs of a
CC malignancy or infection or have an immunodeficiency.
XX

SQ Sequence 127 AA;

Query Match 93.6%; Score 88; DB 23; Length 127;
Best Local Similarity 94.4%; Pred. No. 4.6e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TEVSHHSRRRLRVNMC 18
|||||||
DB 32 TEVSHHSRRRLRVNMC 49

RESULT 15
AAB01452
ID AAB01452 standard; Protein; 130 AA.
XX
AC AAB01452;
XX
DT 20-OCT-2000 (first entry)
XX
DE Rodent Vlc protein sequence.
XX
KW Cutaneous T-cell attracting chemokine; CTRACK; skin; cell movement;
KW migration; vasoactive intestinal contractor; Vlc; GPR2; agonist;
KW antagonist; antibody; immunological condition; mutein.
XX
OS Mus sp.
XX
PN WO200038713-A1.
XX
PD 06-JUL-2000.
XX
PF 23-DEC-1999; 99WO-US30819.
XX
PR 24-DEC-1998; 98US-0113858.
XX
PR 27-MAY-1999; 99US-0322580.
XX
PA (SCHE) SCHERING CORP.
XX
PI Wang W, Oldham ER, Soto H, Lui Y, Hudak SA, Homey B, Morales JM;
PI Kellermann S, McEvoy LM, Zlotnik A;
XX
DR WPI: 2000-465633/40.
XX
DR N-PSDB: AAA47547.
XX
PT Modulating cell movement within the skin, useful for treating
PT immunological skin conditions or diseases comprises administering T
PT cell-attracting chemokine or vasoactive intestinal contractor chemokine
PT agonists or antagonists
XX
PS Example 3; Page 72; 79pp; English.
XX
CC Modulating movement of a cell within or to the skin of a mammal can
CC be achieved by administering an antagonist or agonist of cutaneous T
CC cell-attracting chemokine (CTRACK) or vasoactive intestinal contractor

CC (VIC) chemokine. The antagonist is selected from a mutain of natural
CC CTACK or VIC, an antibody which neutralises CTACK or VIC or an
CC antibody which block GPR2 ligand binding. The CTACK or VIC agonists
CC or antagonists are useful for treating medical conditions or diseases
CC associated with immunological conditions of the skin.

XX
SQ Sequence 130 AA;

Query Match 79.8%; Score 75; DB 21; Length 130;
Best Local Similarity 77.8%; Pred. No. 7.1e-05;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 TEVSHHISRLLERYVMC 18
| | | | | : | | | | | : |
DB 32 TEVSHHVSGRLLERYVSSC 49

Search completed: January 14, 2003, 18:13:15
Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

GN C35D10.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RU Science 287:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Fulton L.;
 RA "The sequence of C. elegans cosmid C35D10."
 RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX Waterston R.;
 RA "Direct Submission";
 RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U21324; AAK68226.2; -.
 KW Hypothetical protein.
 SO SEQUENCE 583 AA; 67741 MW; 24F5C265B7A4C082 CRC64;

Query Match 50.5%; Score 49; DB 5; Length 583;
 Best Local Similarity 60.0%; Pred. No. 9.6;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 NVCHRRKHHG 15
 DB 125 NMCHEREHNG 134

RESULT 3

OS88N4 PRELIMINARY; PRT; 633 AA.

AC OS88N4;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE Putative Ser/Thr protein kinase.

GN AT2G46850.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustoids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
 RA Shen M., Romling C.M., Fraser C.M., Somerville C.R., Venter J.C.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;
 RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV. COLUMBIA;
 RA Town C.D., Kaul S.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC005310; AAM15020.1; -.

KW kinase.

SO SEQUENCE 633 AA; 70166 MW; DB7C35A9FC66FD7 CRC64;

Query Match 46.9%; Score 45.5; DB 10; Length 633;

Best Local Similarity 52.6%; Pred. No. 38;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
 OY 1 KNGK---GNVCHRRKHHG 16
 DB 264 KDGKELYGDKCKIKHHNG 282

RESULT 4

O43169 PRELIMINARY; PRT; 377 AA.

AC O43169;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE Alcohol dehydrogenase (EC 1.1.1.1).

OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI_TaxId=4113;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=PISTIL;
 RA Van Elvik G.J., Rulter R.K., Reijnen W.H., Van Herpen M.M.A.,
 RA Schrauwen J.A.M., Willems G.J.;

RT "An alcohol dehydrogenase-like gene is pistil-specifically expressed
 RT in Solanum tuberosum.";
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

CC 1-COFACITOR: ZINC (BY SIMILARITY).
 DR EMBL; X92179; CAA63093.1; -.

DR HSP; P11766; ITH.

DR InterPro; IPR002328; ADH_zinc.

DR InterPro; IPR002085; Adh_zn_family.

DR InterPro; IPR000205; NAD_binding.

DR Pfam; PF00107; adh_zinc; 1.

DR PROSITE; PS00059; ADH_ZINC; 1.

KW Oxidoreductase; Zinc.
 SO SEQUENCE 377 AA; 40928 MW; C9B0EDBF2A7D1C0 CRC64;

Query Match 46.4%; Score 45; DB 10; Length 377;
 Best Local Similarity 57.1%; Pred. No. 28;
 Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

OY 1 KNGKGNVCHRRKH 14
 DB 106 KSGKSNLCH-KYH 117

RESULT 5

O05874 PRELIMINARY; PRT; 427 AA.

AC O05874;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DE Hypothetical 48.4 kDa protein (Linoleoyl-CoA desaturase,
 DE putative).

GN DESA3 OR RV3229C OR MTCY20B11.04C OR MT3326.

OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxId=1773;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skellon S., Squares S., Squares R.,
RA Sullivan J.E., Taylor K., Whitehead S., Barrell B.G.,
RT "Deciphering the Biology of Mycobacterium tuberculosis from the
RT complete genome sequence." ;
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC SRRAIN-CDC 1551 / OSMKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Esmolaeva M.D., Salzberg S.
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.
RA Bishel W. ;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains." ;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
RL EMBL: Z95121; CAB08330.1. "-
RL EMBL: AE007144; AKK47669.1. "-
DR TIGR: MW3326; "-
DR TubercuList; RV3229C; "-
DR InterPro; IPR001235; FA_desaturase.
DR Pfam; PF00487; FA_desaturase; 1.
DR ProDom; PD001081; FA_desaturase; 1.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 427 AA; 48443 MW; 56341e7f1b393a65 CMC6;

QY	5	GNVCHRRKKHH	14	11:11:11	46.4%	Score 45	DB 16	Length 427	
		11:11:11			Best Local Similarity	60.0%			
DB	300	GNLCHQIEHH	309		Matches 6	Conservative 3	Mismatches 1	Indels 0	Gaps 0

```

RESULT 6
Q9FFP8
ID Q9FFP8 . PRELIMINARY; PRT; 73 AA.
AC Q9FFP8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Genomic DNA, chromosome 5, p1 clone:MBK5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsidae.
OX NCBI_TaxId=3702;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=COLUMBIN;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneo T., Asakami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 mb regions covered by twenty physically assigned
RT pl clones."
RL DNA Res. 4:215-230(1997).
DR EMBL: AB005234; BABI0459.1; -.
DR HSSP; P20230; 1GPT.
DR InterPro: IPR002118; Gamma-thionin.
DR InterPro: IPR003614; Knoc1.
DR Pfam: PF00304; Gamma-thionin; 1.
DR ProDom: PD002594; Gamma-thionin; 1.
DR SMART; SM00505; Knoc1.1.
SQ SEQUENCE 73 AA; 8388 MW; 161507EA9C2A7509 CRC64;

```

Query Match	45.4%	Score 44;	DB 10;	Length 73;
Best Local Similarity	50.0%;	Pred. No. 8.2;		
Matches	7;	Conservative	1;	Mismatches 6;
				Indels 0;
				Gaps 0;
QY	3	CKGNVCHRRKKHGGK	16	
	:			

Db	24	GEORCSKSHRK	37
RESULT	7		
ID	087597	PRELIMINARY;	PRT; 170 AA.
AC	087597;		
DT	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	Hypothetical 18.6 kDa protein.		
OS	Streptomyces coelicolor.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces		
OX	NCBI_TaxId=1902;		
..RN	[1]		
..RN	SEQUENCE FROM N.A.		
RP	STRAIN-A3(2);		
RC			
RX	MEDLINE=98440439; PubMed=9765579;		
RA	Nagy I., Tamura T., Vanderleyden J., Baumeister W., De Mot R.;		
RT	"The 20S proteasome of Streptomyces coelicolor.";		
RL	J. Bacteriol. 180:5448-5453(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-A3(2);		
RA	De Mot R.;		
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF086832; AAC64277.1; "		
DR	InterPro: IPR004211; Endonuclease_7.		
DR	Pfam: PF02945; endonuclease_7; 1.		
KW	Hypothetical protein		
SO	SEQUENCE 170 AA; 18556 MW; 413EPD8E0C830F77 CRC64;		

Query Match	45.4%	Score 44	DB 2	Length 170
Best Local Similarity	53.3%	Pred. No. 19		
Matches	8	Conservative	2	Mismatches 5
				Indels 0
				Gaps 0
OY	1	KNKGNCVCHRRKHHG	15	
db	71	KAAKGRVGHILKRRHG	85	

```

RESULT 8
ORJ59
ID ORJ59 PRELIMINARY: PRT: 197 AA.
AC ORJ59:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SC01645.
GN SC01645 OR SC141.28C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID:1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2).
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2):
RA Cerdeno A.M., Partholl J., Barrell B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2):
MEDLINE=97000351: Pubmed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Klnasht H., Hopwood D.A.;
RN "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
```

RL Mol. Microbiol. 21:77-96(1996).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Godle A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).
 RL Nature 417:141-147(2002).
 DR EMBL: AL132648; CAB59498.1;
 DR InterPro: IPR004211; Endonuclease_7.
 DR Pfam: PF02945; endonuclease_7; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 197 AA; 21640 MW; AFEF7A2E19C4A451 CRC64;

Query Match 45.4%; Score 44; DB 16; Length 197;
 Best Local Similarity 53.3%; Pred. No. 21;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNGKGVNCHRRKHG 15
 Db 98 KAKGRVGHKRYHG 112

RESULT 9
 O9CL67 PRELIMINARY; PRT; 303 AA.
 AC O9CL67;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein PM1375.
 GN PM1375.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OC NCBI_TaxID=747;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-PM70;
 RX MEDLINE-21145866; PubMed-11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.,
 RT Complete genome sequence of Pasteurella multocida pm70.
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 DR EMBL: AE006175; AAK03459.1;
 DR InterPro: IPR000600; ROK_family.
 DR Pfam: PF00480; ROK; 1.
 DR PROSITE: PS01125; ROK; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 303 AA; 32405 MW; A4119B83FE7E CRC64;

Query Match 45.4%; Score 44; DB 16; Length 303;
 Best Local Similarity 61.5%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 GKGNVCHRRKHG 15
 Db 135 GSGIVHGRPHG 147

RESULT 10
 O93968 PRELIMINARY; PRT; 317 AA.
 AC O93968;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE T01H8.2 protein.
 GN T01H8.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN
 RP SEQUENCE FROM N.A.
 RA Lennard N.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Investigating biology".
 RT Science 282:2012-2018(1998).
 RL EMBL: 280219; CAB02298.1;
 DR HSSP: P35520; 1JBQ.
 DR InterPro: IPR001926; B6_enzyme_beta.
 DR InterPro: IPR000634; S/T_dehydrtse.
 DR Pfam: PF00291; PALP.1.
 DR PROSITE: PS00165; DEHYDRATASE_SER_THR.1.
 SQ SEQUENCE 317 AA; 34688 MW; OFDDE6C7479389E3 CRC64;

Query Match 45.4%; Score 44; DB 5; Length 317;
 Best Local Similarity 43.8%; Pred. No. 34;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 KNGKGVNCHRRKHG 16
 Db 70 KNAKGIAHSSGNHG 85

RESULT 11
 O9NXL8 PRELIMINARY; PRT; 358 AA.
 AC O9NXL8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE CDNA FLJ20171 f1s, clone COL09761.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE-COLON;
 RA Kawabata A., Hiki J., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEO human cDNA sequencing project".
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK000178; BAA90992.1;
 DR InterPro: IPR000504; RNA_rec_mot.
 DR InterPro: IPR003955; RRM_2.
 DR Pfam: PF00076; rrm; 3.
 DR SMART: SM00360; RRM; 2.
 DR SMART: SM00362; RRM_2; 1.
 DR PROSITE: PS50102; RRM; 2.
 SQ SEQUENCE 358 AA; 40109 MW; 69A0A916F93EF768 CRC64;

Query Match 45.4%; Score 44; DB 4; Length 358;
 Best Local Similarity 66.7%; Pred. No. 38;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 8 CHRRKHG 16
 Db 348 CHRRKHG 356

RESULT 12
 095TH1 PRELIMINARY: PRT: 558 AA.
 ID 095TH1
 AC 095TH1
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Stress-induced protein still-like protein.
 GN TAC9.240 OR AT4G12400.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Beyer M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
 RA Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Mannhaupt G.,
 RA Schueller C.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
 RA Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL080318; CAB45987.1; -
 DR EMBL: AL161534; CAB78283.1; -
 DR HSSP: P53041; 1A17.
 DR InterPro: IPR01440; TPR.
 DR Pfam: PF00515; TPR; 9.
 DR SMART: SM00028; TPR; 9.
 SO SEQUENCE 558 AA; 63706 MW; 47010D35F0F98DB9 CRC64;

Query Match 45.4%; Score 44; DB 10; Length 558;
 Best Local Similarity 43.8%; Pred. No. 59;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 KNGKNVCHRRKHGK 16
 DB 233 EKGSNVAKKKDFGR 248

RESULT 13
 044728 PRELIMINARY: PRT: 1017 AA.
 ID 044728
 AC 044728
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical 113.7 kDa protein.
 GN F49D11.9.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peioderinae; Caenorhabditis.
 NC NCBI_Taxid=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BRISTOL N2;
 RC MEDLINE-99069613; PubMed-9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;
 RA Hinds, K., Graves T., Duckels G., Ozersky P.;
 RT "The sequence of C. elegans cosmid F49D11.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BRISTOL N2;
 RC Waterston R.;
 RT "Direct Submission.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
 DR EMBL: AF039711; AAB96714.1; -
 DR InterPro: IPR01680; WD40.
 DR Pfam: PF00400; WD40; 2.
 KW Hypothetical protein; Repeat; WD repeat.
 SO SEQUENCE 1017 AA; 113651 MW; 287C632DBF9CCAB CRC64;

Query Match 45.4%; Score 44; DB 5; Length 1017;
 Best Local Similarity 53.8%; Pred. No. 1e+02;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNGKNVCHRRKH 13
 DB 926 KTAGIICHRRPH 938

RESULT 14
 041622 PRELIMINARY: PRT: 275 AA.
 ID 041622
 AC 041622
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Complex specificity lectin precursor.
 OS Tulipa sp.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae; Tulipa.
 NC NCBI_Taxid=45423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV APBLDOOR;
 RC MEDLINE-96195647; PubMed-8612611;
 RA Van Damme E.J., Brike F., Winter H.C., Van Leuven F., Goldstein I.J.,
 RA Peumans W.J.;
 RT "Molecular cloning of two different mannose-binding lectins from tulip
 RT bulbs.";
 RL Eur. J. Biochem. 236:419-427(1996).
 DR EMBL: U23041; AAC49384.1; -
 DR HSSP: P30617; LIPC.
 DR InterPro: IPR01480; B_lectin.
 DR Pfam: PF01453; Agglutinin; 2.
 DR SMART: SM00108; B_lectin; 2.
 KW Lectin; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 275
 SO SEQUENCE 275 AA; 29726 MW; 05BB3C45980B18F CRC64;

Query Match 44.8%; Score 43.5; DB 10; Length 275;
 Best Local Similarity 56.2%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 1 KNGKNVCH-RKHHG 15
 DB 202 ENKKNVCHVRDLHRG 217

RESULT 15
 09ASJ3 PRELIMINARY: PRT: 294 AA.
 ID 09ASJ3
 AC 09ASJ3
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE P0439B06.19 protein.
GN P0439B06.19.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriobotryaceae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0439B06." to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002882; BAB39884.1; -
DR InterPro: IPR001092; HLH_basic.
DR SMART: SM00353; HLH; 1.
SO SEQUENCE 294 AA; 30541 MW; C57B8538C964F3A CRC64;

Query Match 44.3%; Score 43; DB 10; Length 294;
Best Local Similarity 53.8%; Pred. No. 46;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 GCGNVCHRRKHHG 15
DB 114 GGNVCHRRLOPHG 126

Search completed: January 14, 2003, 18:15:30
Job time : 10.0889 secs

Rülin, X.; Kaul, S.; Rounsfley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,

enus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-766, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: A84908
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-633 <STO>
 A:Cross-references: GB:AE002093; NID:96598416; PIDN:CAAC34215.2; GSPDB:GN00139
 C:Genetics:
 A:Gene: AT2g46850; F19D11.13
 A:Map position: 2
 A:Introns: 261/1; 289/1

Query Match 46.9%; Score 45.5; DB 2; Length 633;
 Best Local Similarity 52.6%; Pred. No. 31;
 Matches 10; Conservative 3; Mismatches 3; Indels 3; Gaps 1;
 OY 1 KNGK---GNVCHRRKKHKK 16
 DB 264 KDGKELYGDKCKIKKHNGK 282

RESULT 3
 T07179
 Probable alcohol dehydrogenase (EC 1.1.1.1) (clone CP67) - potato
 C:Species: Solanum tuberosum (potato)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T07179
 R:van Eldik, G.J.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z15978
 A:Accession: T07179
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-377 <VAN>
 A:Cross-references: EMBL:X92179; NID:q1039354; PIDN:CAA63093.1; PID:q1039355
 A:Experimental source: pistill
 C:Function:
 A:Description: catalyzes the oxidation of primary and secondary alcohols to aldehydes an
 A:Pathway: alcohol degradation
 C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
 C:Keywords: alcohol metabolism; metalloprotein; NAD; oxidoreductase

Query Match 46.4%; Score 45; DB 2; Length 377;
 Best Local Similarity 57.1%; Pred. No. 24;
 Matches 8; Conservative 3; Mismatches 1; Indels 2; Gaps 1;
 OY 1 KNGKGNVCHRRKKH 14
 DB 106 KSGKSNLCH--KYH 117

RESULT 4
 G70590
 Probable dea3 protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: G70590
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: G70590
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-427 <COL>
 A:Cross-references: GB:Z95121; GB:AL123456; NID:93261742; PIDN:CAB08330.1; PID:ej14467;
 A:Experimental source: strain H37RV
 C:Genetics:

A:Gene: dea3
 Query Match 46.4%; Score 45; DB 2; Length 427;
 Best Local Similarity 60.0%; Pred. No. 27;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 5 GNVCHRRKKH 14
 DB 300 GNLCHQIEHH 309

RESULT 5
 T24337
 threonine ammonia-lyase (EC 4.3.1.19) T01H8.2 [similarity] - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jun-2002
 C:Accession: T24337
 R:Lennard, N.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19877
 A:Accession: T24337
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-317 <MIT>
 A:Cross-references: EMBL:Z80219; PIDN:CAB02298.1; GSPDB:GN00019; CESP:T01H8.2
 A:Experimental source: clone T01H8
 C:Genetics:
 A:Gene: CESP:T01H8.2
 A:Map position: 1
 A:Introns: 75/3; 131/2; 159/3; 203/1; 259/3; 276/3
 C:Superfamily: threonine dehydratase
 C:Keywords: ammonia-lyase; carbon-nitrogen lyase; carbon-oxygen lyase; hydro-lyase; p
 F.56/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 45.4%; Score 44; DB 2; Length 317;
 Best Local Similarity 43.8%; Pred. No. 30;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 OY 1 KNGKGNVCHRRKKHKK 16
 DB 70 KNAKGMHSSNHHQ 85

RESULT 6
 T48150
 stress-induced protein still-like protein - *Arabidopsis thaliana*
 N:Alternate names: protein T4C9.240
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
 C:Accession: T48150
 R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z24485
 A:Accession: T48150
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-558 <BEV>
 A:Cross-references: EMBL:AL080318
 A:Experimental source: cultivar Columbia; BAC clone T4C9
 C:Genetics:
 A:Map position: 4
 A:Introns: 279/3; 376/1; 479/2; 521/3
 A>Note: T4C9.240
 C:Superfamily: unassigned tetrairicopeptide repeat proteins; tetrairicopeptide repeat

Query Match 45.4%; Score 44; DB 2; Length 558;
 Best Local Similarity 43.8%; Pred. No. 47;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 OY 1 KNGKGNVCHRRKKHKK 16
 DB 233 EKGGNVAVKKKDFGR 248

RESULT 7
T37201
hypothetical protein F49D11.9 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C/Accession: T37201
R/Links: K.; Graves, T.; Duckels, G.; Ozersky, P.
submitted to the EMBL Data Library, December 1997
A/Description: The sequence of *C. elegans* cosmid F49D11.
A/Reference number: 221632
A/Accession: T37201
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1017 <HIN>
A/Cross-references: EMBL:AF039711; PIDN:AA96714.1
C/Genetics:
A/Map position: 1
A/Intons: 177/2; 212/2; 241/1; 277/2; 353/2; 457/1; 564/1; 598/2; 820/3; 929/3
A/Note: F49D11.9

Query Match
Best Local Similarity 45.4%; Score 44; DB 2; Length 1017;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 KNGKNVCHRRKH 13
DB 926 KTAGKICHRPAH 938

RESULT 8
S62647
mannose-binding lectin I.1 precursor - *Tulipa* sp.
C/Species: *Tulipa* sp. (tulip)
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 10-Dec-1999
C/Accession: S62647; S62644
R/Van Damme, E.T.M.; Brike, F.; Winter, H.C.; Van Leuven, F.; Goldstein, I.J.; Peumans, Eur. J. Biochem. 236, 419-427, 1996
A/Title: Molecular cloning of two different mannose-binding lectins from tulip bulbs.
A/Reference number: S62644; MID:96195647; PMID:8612611
A/Accession: S62647
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-275 <VAN>
A/Cross-references: EMBL:U23041; NID:G1141760; PIDN:AA949384.1; PID:G1141761
A/Experimental source: cultivar Apeldoorn
A/Accession: S62644
A/Molecule type: protein
A/Residues: 46-48, 'X', 50-61; 144-163; 169-173; 'D', 175-182, 'X', 184-187 <DAM>
C/Superfamily: taro globulin G1
F:1-19/Domain1: signal sequence #status predicted <SIG>
F:20-275/Product: mannose-binding lectin I.1 #status predicted <MAT>

Query Match
Best Local Similarity 44.8%; Score 43.5; DB 2; Length 275;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 1 KNGKNVCH-RRKHG 15
DB 202 ENCKENCHVRLDHRG 217

RESULT 9
D89101
protein F25E5.8 [imported] - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C/Accession: D89101
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A/Reference number: A75000; MID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A/Accession: D89101
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-235 <STO>
A/Cross-references: GB:chr.V; PIDN:AA927331.1; PID:G3335238; GSPDB:GN00023; CESP:F25E
A/Note: weak similarity to 'POU' transcription factors
C/Genetics:
A/Map position: 5
A/Intons: 5

Query Match
Best Local Similarity 43.3%; Score 42; DB 2; Length 235;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 NGKGNVCHRRKHG 15
DB 88 SGSSSLKEKKHKG 101

RESULT 10
G81417
cytochrome-c peroxidase (EC 1.11.1.5) Cj0020c [similarity] - *Campylobacter jejuni* (str. C/Species: *Campylobacter jejuni*)
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C/Accession: G81417
R/Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chli C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanyile, A.; Whitehead, S.; Bar Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals
A/Reference number: A81250; MID:20150912; PMID:10688204
A/Accession: G81417
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-304 <PAR>
A/Cross-references: GB:AL139074; GB:AL111168; NID:96967505; PIDN:CAB72513.1; PID:9696
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/KeyWords: Pseudomonas cytochrome-c peroxidase; Pseudomonas cytochrome-c peroxidase
C/Superfamily: Pseudomonas cytochrome-c peroxidase; Pseudomonas cytochrome-c peroxidase
F:56/58/Binding site: heme (Cys) (covalent) (low potential) #status predicted
F:60/259/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:199/202/Binding site: heme (Cys) (covalent) (high potential) #status predicted
F:203/273/Binding site: heme iron (His, Met) (axial ligands) (high potential) #status

Query Match
Best Local Similarity 43.3%; Score 42; DB 2; Length 304;
Matches 9; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

OY 1 KNGK--GNVCHRRKHG 15
DB 50 KDGKVSCTCHRLDHRG 66

RESULT 11
T47146
hypothetical protein DKFZp761C169.1 - human (fragment)
C/Species: *Homo sapiens* (man)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T47146
R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A/Reference number: 224379
A/Accession: T47146
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-492 <AAA>
A/Cross-references: EMBL:AL161991
A/Experimental source: adult amygdala; clone DKFZp761C169
C/Genetics:
A/Note: DKFZp761C169.1

Query Match	43.3%	Score 42	DB 2	Length 492
Best Local Similarity	46.7%	Prod. No. 85		
Matches	7	Conservative	3	Mismatches 5
				Indels 0
				Gaps 0
Oy	1	KNGKGNVCHRRKKNHG	15	
	:		:	
Db	97	RNGTENINNRGCTHG	111	

Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:99196666; PMID:9537320
 A:Accession: D70445
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-603 <AEP>
 A:Cross-references: GB:AE000750; NID:g29383999; PIDN:AAC07548.1; PID:g2984003; GB:AE000654
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: asps
 C:Function:
 A:Description: activates amino acid and transfers it to specific tRNA molecule
 A:Pathway: protein biosynthesis
 C:Superfamily: lysine-tRNA ligase
 C:Keywords: aminocyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match	Similarity	43.3%	Score 42;	DB 2;	Length 603;
Best Local	Similarity 85.7%		Pred. No. 1e+02;		
Matches	6; Conservative	1;	Mismatches	0;	Indels 0; Gaps 0;

C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: J33528
R:Wamsley, P., Tytman, B.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid F58E1.
A:Reference number: 221365
A:Accession: J33528
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-634 <MAN>
A:Cross-references: EMBL:AF098995; PIDN:AAC67478.1; GSPDB:GN00020; CESP:F58E1.6
A:Experimental source: strain Bristol N2; clone F58E1
C:Genetics:
A:Gene: CESP:F58E1.6
A:Map position: 2
A:Introns: 66/1; 155/1; 199/3; 308/2; 399/3; 447/3; 477/3; 533/3

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:40:55 ; Search time 10.2222 Seconds
(without alignments)
208.566 Million cell updates/sec

Title: US-09-834-794A-4
Perfect score: 97
Sequence: 1 KNGKGNVCHRRKHHGK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	97	100.0	16	20 AAY29095	Human mammary asso
2	97	100.0	104	20 AAY29093	Human mammary asso
3	97	100.0	106	23 AAB47869	Mature placenta-de
4	97	100.0	108	20 AAY11658	Human 5' EST seque
5	97	100.0	127	19 AAW60649	Human DY1C-1 C-C
6	97	100.0	127	20 AAY29092	Human mammary asso
7	97	100.0	127	21 AAB01450	Primate VIC protei
8	97	100.0	127	21 AAB03001	Human growth facto
9	97	100.0	127	22 AAB69175	Human G-protein co
10	97	100.0	127	23 AAB07728	Human mucosae-asso

11	97	100.0	127	23 ABB08268	Human mammary glan
12	97	100.0	127	23 AAB47868	Full length placen
13	97	100.0	147	22 AAB25479	Human protein sequ
14	48.5	50.0	1550	23 ABP30895	Streptococcus poly
15	48.5	50.0	1570	23 ABP29894	Streptococcus poly
16	48.5	50.0	1570	23 ABP30531	Streptococcus poly
17	48.5	50.0	1590	23 ABP28560	Streptococcus poly
18	47.5	49.0	609	22 ABB23539	Novel human diagn
19	45.5	46.9	633	23 ABB92137	Herbically activ
20	44	45.4	558	21 AAY77979	A. thaliana enviro
21	43	44.3	130	21 AAB01452	Rodent VIC protein
22	43	44.3	762	21 AAB90977	N. dentriticans a
23	43	44.3	1156	22 AABG09141	Novel human diagn
24	43	44.3	1156	22 AABG18480	Novel human diagn
25	42	43.3	56	22 AAU66116	Propionibacterium
26	42	43.3	76	22 AAU62426	Propionibacterium
27	42	43.3	329	22 ABB23127	Novel human diagn
28	42	43.3	473	22 AAG68752	Human ATP-depend
29	42	43.3	400	22 AAG78852	Human ORFX ORF3029
30	42	43.3	500	21 AAB43265	Novel human diagn
31	42	43.3	516	22 ABB18191	Protein encoded by
32	42	43.3	525	23 AAB47788	Drosophila melanog
33	42	43.3	129	23 ABB64623	Corn KCP-like prot
34	41.5	42.8	80	23 AAO21288	Human ovarian anti
35	41	42.3	80	23 ABB43203	Propionibacterium
36	41	42.3	82	22 AAU47366	Marine immune asso
37	41	42.3	346	22 AAB36217	Drosophila melanog
38	40.5	41.8	647	22 ABB67093	Human ORFX protein
39	40.5	41.8	672	22 ABB54473	Propionibacterium
40	40	41.2	50	23 ABB02930	Human ORFX protein
41	40	41.2	60	23 ABB11399	Novel human diagn
42	40	41.2	74	22 AAU49312	Human protein SEQ
43	40	41.2	164	22 ABB10158	Human protein
44	40	41.2	164	22 ABB03300	Human protein
45	40	41.2	166	22 AAG92528	Human protein

ALIGNMENTS

RESULT 1
AAY29095
ID AAY29095 standard; peptide: 16 AA.
XX
AC AAY29095;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human mammary associated chemokine (MACK) protein fragment MACK B.
XX
KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
KW Inflammation; infection; mastitis; benign cystitis; hyperplasia;
KW mammary associated chemokine; MACK; epitope.
XX
OS Homo sapiens.
XX
PN MO9936540-A1.
XX
PD 22-JUL-1999.
XX
PF 12-JAN-1999; 99WO-US00651.
XX
PR 09-JUL-1998; 98US-0092155.
XX
PR 20-JAN-1998; 98US-0071899.
XX
PA (CODON) CODON DIAGNOSTICS LLC.
XX
PI Dyster LM, Frustaci JM, Papsidero LD;
XX
DR WPI: 1999-458469/38.
XX
PT A mammary associated chemokine and related polynucleotides, useful
for detection and treatment of breast disease, especially cancer

XX PS Claim 9; Page 67; 76pp; English.

XX CC The invention provides an isolated human chemokine, which is

CC preferentially expressed in breast tissue or detected in breast milk. An

CC antibody that recognizes the novel chemokine, or a chemokine-derived

CC antigenic peptide, can be used to treat breast disease in a patient. A

CC peptide, which binds to a cellular receptor for the chemokine, can also

CC be used to treat breast disease. Antigenic peptides of the chemokine

CC can be used to vaccinate patients against breast disease. The chemokine

CC polynucleotide sequences and the chemokine protein can be detected in

CC samples with primers, probes and antibodies using standard techniques.

CC This is useful for detecting breast disease. Other breast diseases that

CC may be treated or detected with the chemokine and its encoding

CC polynucleotides include inflammations, infections, mastitis, benign

CC cystitis, and benign hyperplasias as well as other malignancies.

CC Sequences AAY29094-96 represent fragments of the human mammary

CC associated chemokine (MACK) protein. These fragments act as antigenic

CC peptide epitopes against which rabbit antisera can be raised.

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 97; DB 20; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.9e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16

DB 1 KNGKGNVCHRRKHHGK 16

RESULT 2

AAY29093

ID AAY29093 standard; Protein; 104 AA.

XX AC AAY29093;

XX DT 29-SEP-1999 (first entry)

XX DE Human mammary associated chemokine (MACK) mature protein sequence.

XX KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;

XX KM Inflammation; infection; mastitis; benign cystitis; hyperplasia;

XX KW mammary associated chemokine; MACK.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT MISC-difference 47 /label= unknown

FT MISC-difference 68 /note= "encoded by NGA"

FT MISC-difference 68 /label= unknown

XX PN WO9936540-A1.

XX PD 22-JUL-1999.

XX PF 12-JAN-1999; 99WO-US00651.

XX PR 09-JUL-1998; 98US-0092155.

XX PR 20-JAN-1998; 98US-0071899.

XX PA (CODON) CODON DIAGNOSTICS LLC.

XX PI Dyster LM, Frustaci JM, Papsidero LD;

XX DR WPI: 1999-458469/38.

XX DR N-PSDB; AAX89389, AAX89390.

XX PT A mammary associated chemokine and related polynucleotides, useful

PT for detection and treatment of breast disease, especially cancer

XX PS Claim 5; Page 66; 76pp; English.

XX CC The invention provides an isolated human chemokine, which is

CC preferentially expressed in breast tissue or detected in breast milk. An

CC antibody that recognizes the novel chemokine, or a chemokine-derived

CC antigenic peptide, can be used to treat breast disease in a patient. A

CC peptide, which binds to a cellular receptor for the chemokine, can also

CC be used to treat breast disease. Antigenic peptides of the chemokine

CC can be used to vaccinate patients against breast disease. The chemokine

CC polynucleotide sequences and the chemokine protein can be detected in

CC samples with primers, probes and antibodies using standard techniques.

CC This is useful for detecting breast disease. Other breast diseases that

CC may be treated or detected with the chemokine and its encoding

CC polynucleotides include inflammations, infections, mastitis, benign

CC cystitis, and benign hyperplasias as well as other malignancies. The

CC present sequence represents the human mammary associated chemokine (MACK)

CC mature protein sequence.

XX SQ Sequence 104 AA;

Query Match 100.0%; Score 97; DB 20; Length 104;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16

DB 69 KNGKGNVCHRRKHHGK 84

RESULT 3

AA87869

ID AAB47869 standard; Protein; 106 AA.

XX AC AAB47869;

XX DT 15-APR-2002 (first entry)

XX DE Mature placenta-derived C-C chemokine.

XX KW Chemokine; placenta-derived; C-C chemokine; PLACC; inflammation;

XX KM immune response; cell proliferation; asthma; malignancy; infection;

XX KW adult respiratory distress syndrome; rheumatoid arthritis;

XX KW lupus erythematosus; psoriasis; osteoarthritis; glomerulonephritis;

XX KW osteoporosis; dermatomyositis; polymyositis; Addison's disease;

XX KW Graves Disease; Crohn's Disease; Irritable bowel syndrome;

XX KW atrophic gastritis; graft versus host disease; myasthenia gravis;

XX KW multiple sclerosis; autoimmune thyroiditis; ulcerative colitis;

XX KW atherosclerosis; pancreatitis.

XX OS Homo sapiens.

XX PN WO200192301-A2.

XX PD 06-DEC-2001.

XX PF 23-MAY-2001; 2001WO-US16599.

XX PR 26-MAY-2000; 2000US-207578P.

XX PA (INDV) UNIV INDIANA ADVANCED RES & TECHNOLOGY.

XX PI Hromas RA;

XX DR WPI: 2002-154522/20.

XX DR N-PSDB; AAI72313.

XX PT New placenta-derived human C-C chemokine, useful for modulating an

PT inflammatory or immune response, particularly for treating asthma,

PT psoriasis, osteoarthritis, Graves Disease, Crohn's Disease or graft

PT versus host disease

XX PS Disclosure; Fig 1; 97pp; English.

XX This sequence shows a chemokine protein, which is a placenta-derived
 CC C-C chemokine (PLACC). The PLACC polypeptide or the chemokine-encoding
 CC polynucleotide, are useful for modulating an inflammatory/immune
 CC response, modulating the activity of PLACC and screening for modulators,
 CC and in stimulating cell proliferation. Particularly these are useful
 CC for treating asthma, adult respiratory distress syndrome, rheumatoid
 CC arthritis, lupus erythematosus, psoriasis, osteoarthritis,
 CC glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
 CC Addison's Disease, Graves Disease, Crohn's Disease, irritable bowel
 CC syndrome, atrophic gastritis, graft versus host disease, myasthenia
 CC graves, multiple sclerosis, autoimmune thyroiditis, ulcerative colitis,
 CC atherosclerosis or pancreatitis. Increasing the PLACC activity in the
 CC subject is useful where the subject may be exhibiting signs of a
 CC malignancy or infection or have an immunodeficiency.

SO Sequence 106 AA:

Query Match 100.0%; Score 97; DB 23; Length 106;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
 Db 71 KNGKGNVCHRRKHHGK 86

RESULT 4
 AAY11658
 ID AAY11658 standard; Protein: 108 AA.

XX AAY11658:
 DT 16-JUN-1999 (first entry)
 XX
 DE Human 5' EST secreted protein SEQ ID NO:310.
 XX
 XX Human: secreted protein; EST; expressed sequence tag; diagnosis;
 KM forensic; gene therapy; chromosome mapping; signal peptide;
 KM upstream regulatory sequence; cytokine activity; cell proliferation;
 KM differentiation; haemopoiesis regulation; tissue growth regulation;
 KM reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KM thrombolytic; anti-inflammatory; tumour inhibition.

XX Homo sapiens.
 OS
 XX MO9906439-A2.
 PN
 XX 11-FEB-1999.
 PD
 XX 31-JUL-1998; 98MO-1801233.
 PF
 XX 01-AUG-1997; 97US-0904468.
 PR
 XX (GEST) GENSET.
 PA
 XX Duclelet A, Dumas Milne Edwards J, Lacroix B;
 PI
 XX WPI: 1999-153700/13.
 DR
 XX N-PSDB: AAX40376.
 DR
 XX New nucleic acids encoding human secreted proteins - obtained from
 PT cDNA libraries derived from liver, lung, large intestine, colon,
 PT thyroid and pancreas tissue
 XX
 PS Claim 27; Page 384-385; 398pp; English.

XX AAX40251 to AAX40397 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY11533 to
 CC AAY11679, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The

CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haemopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

SO Sequence 108 AA:

Query Match 100.0%; Score 97; DB 20; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHHGK 16
 Db 92 KNGKGNVCHRRKHHGK 107

RESULT 5
 AAM60649
 ID AAM60649 standard; Protein: 127 AA.

XX AAM60649:
 DT 12-OCT-1998 (first entry)
 XX
 DE Human DVlc-1 C-C chemokine.
 XX
 XX DVlc-1; DNAX Vlc-1; C-C chemokine; cytokine; human; immune system;
 KM cancer; cell proliferation; therapy; diagnosis.
 KM
 XX Homo sapiens.
 OS
 XX Key location/Qualifiers
 FH Peptide 1..22
 FT /label= Sig-peptide
 FT 23..127
 FT Protein /label= Mat.protein
 FT /note= "Claim 1"

XX MO9823750-A2.
 PN
 XX 04-JUN-1998.
 PD
 XX 26-NOV-1997; 97WO-US21092.
 PF
 XX 05-DEC-1996; 96US-0761071.
 PR
 XX 27-NOV-1996; 96US-0031805.
 PR
 XX (SCHE) SCHERING CORP.
 PA
 XX Hedrick JA, Morales J, Vicari A, Zlotnik A;
 PI
 XX WPI: 1998-322730/28.
 DR
 XX N-PSDB: AAV36291.
 DR
 XX DVlc-1 and DGMCC chemokines - useful for developing products for
 PT treating abnormal physiology or development, e.g. cancerous or
 PT degenerative conditions
 XX
 PS Claim 1; Page 59-60; 71pp; English.

XX This polypeptide comprises human DNAX Vlc-1 (DVlc-1), a novel C-C
 CC chemokine, the mature portion of which is claimed. The amino acid
 CC sequence was deduced from a cDNA clone (see AAV36291). An alternative
 CC longer transcript (see AAM60652) for human DVlc-1 is also disclosed.
 CC Also claimed is novel human DNAX Groin wound expressed CC chemokine
 CC (DGMCC) (see AAM60649) mature protein, as well as expression vectors

CC and host cells. DVC-1 and DGMCC play a role in the regulation or
 CC development of neuronal or hematopoietic cells, e.g. lymphoid
 CC cells, which affect immunological responses. They can be used in
 CC the treatment of conditions associated with abnormal physiology or
 CC development, including abnormal proliferation, e.g. cancerous
 CC conditions or degenerative conditions. Abnormal proliferation,
 CC regeneration, degeneration, and atrophy may be modulated by
 CC appropriate therapeutic treatment using products of the invention.
 CC The products can also be used for detection, diagnosis and drug
 CC screening.
 CC
 SQ Sequence 127 AA;
 Query Match 100.0%; Score 97; DB 19; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KNGKGNVCHRRKHHGK 16
 ||||||||||||
 DB 92 KNGKGNVCHRRKHHGK 107
 RESULT 6
 AAY29092
 ID AAY29092 standard; Protein; 127 AA.
 XX
 AC AAY29092;
 XX
 DT 29-SEP-1999 (first entry)
 XX
 DE Human mammary associated chemokine (MACK) protein.
 XX
 KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
 KW Inflammation; infection; mastitis; benign cystitis; hyperplasia;
 KW mammary associated chemokine; MACK.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23 /note= "signal peptide"
 FT Protein 24..127 /note= "mature protein (AAY29093)"
 FT Region 32..49 /note= "antigenic peptide epitope MACK A (AAY29094)"
 FT Misc-difference 70 /label= "unknown
 /note= "encoded by NGA"
 FT Misc-difference 91 /label= "unknown
 /note= "encoded by AAN"
 FT Region 92..107 /note= "antigenic peptide epitope MACK B (AAY29095)"
 FT Region 109..127 /note= "antigenic peptide epitope MACK C (AAY29096)"
 FT
 XX WO936540-A1.
 XX
 XX 22-JUL-1999.
 XX
 XX 12-JAN-1999; 99WO-US00651.
 XX
 XX 09-JUL-1998; 98US-0092155.
 XX 20-JAN-1998; 98US-0071899.
 XX
 XX (CODON) CODON DIAGNOSTICS LLC.
 XX
 XX Dyster LM, Frustaci JM, Papsidero LD;
 XX
 XX WPI; 1999-458469/38.
 XX N-PSDB; AAX89389, AAX89390.
 XX
 PT A mammary associated chemokine and related polynucleotides, useful

PT for detection and treatment of breast disease, especially cancer
 XX Claim 4; Page 48; 76pp; English.
 XX
 CC The invention provides an isolated human chemokine, which is
 CC preferentially expressed in breast tissue or detected in breast milk. An
 CC antibody that recognizes the novel chemokine, or a chemokine-derived
 CC antigenic peptide, can be used to treat breast disease in a patient. A
 CC peptide, which binds to a cellular receptor for the chemokine, can also
 CC be used to treat breast disease. Antigenic peptides of the chemokine can
 CC be used to vaccinate patients against breast disease. The chemokine
 CC polynucleotide sequences and the chemokine protein can be detected in
 CC samples with primers, probes and antibodies using standard techniques.
 CC This is useful for detecting breast disease. Other breast diseases that
 CC may be treated or detected with the chemokine and its encoding
 CC polynucleotides include inflammations, infections, mastitis, benign
 CC cystitis, and benign hyperplasias as well as other malignancies. The
 CC present sequence represents the amino acid sequence of the human mammary
 CC associated chemokine (MACK) protein.
 CC
 SQ Sequence 127 AA;
 Query Match 100.0%; Score 97; DB 20; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KNGKGNVCHRRKHHGK 16
 ||||||||||||
 DB 92 KNGKGNVCHRRKHHGK 107
 RESULT 7
 AAB01450
 ID AAB01450 standard; Protein; 127 AA.
 XX
 AC AAB01450;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Primate Vlc protein sequence.
 XX
 KW Cutaneous T-cell attracting chemokine; CTRAC; skin; cell movement;
 KW migration; vasoactive intestinal contractor; Vlc; GPR; agonist;
 KW antagonist; antibody; immunological condition; mutein.
 XX
 OS Homo sapiens.
 XX
 PN WO200038713-A1.
 XX
 XX 06-JUL-2000.
 XX
 XX 23-DEC-1999; 99WO-US30819.
 XX
 XX 24-DEC-1998; 98US-0113858.
 XX 27-MAY-1999; 99US-0322580.
 XX
 XX (SCHE) SCHERING CORP.
 XX
 XX Wang W, Oldham ER, Soto H, Lui Y, Hudak SA, Homey B, Morales JM;
 XX Kellermann S, McEvoy LM, Zlotnik A;
 XX
 XX WPI; 2000-465633/40.
 XX N-PSDB; AAA47545.
 XX
 PT Modulating cell movement within the skin, useful for treating
 PT immunological skin conditions or diseases comprises administering T
 PT cell-attracting chemokine or vasoactive intestinal contractor chemokine
 PT agonists or antagonists
 XX
 XX Example 3; Page 69; 79pp; English.
 XX
 CC Modulating movement of a cell within or to the skin of a mammal can
 CC be achieved by administering an antagonist or agonist of cutaneous T

CC cell-attracting chemokine (CTACK) or vasoactive intestinal contractor
 CC (VIC) chemokine. The antagonist is selected from a mutant of natural
 CC CTACK or VIC, an antibody which neutralises CTACK or VIC or an
 CC antibody which block GPR2 ligand binding. The CTACK or VIC agonists
 CC or antagonists are useful for treating medical conditions or diseases
 CC associated with immunological conditions of the skin.

XX Sequence 127 AA;

Query Match 100.0%; Score 97; DB 21; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHGK 16
 |||||
 DB 92 KNGKGNVCHRRKHGK 107

RESULT 8

AAB03001 standard; Protein: 127 AA.

XX AAB03001;

XX 25-SEP-2000 (first entry)

XX Human growth factor related molecule GFRP-2.

XX Human GFRP-2: growth factor related molecule; breast tissue;
 KW breast tumour; CC chemokine; hTECK homologue; developmental disorder;
 KW cell proliferative disorder; immune disorder; reproductive disorder;
 KW cardiovascular disorder; bacterial infection; viral; fungal; parasitic;
 KW cancer; allergy; asthma; arteriosclerosis; therapy; diagnosis.

OS Homo sapiens.

XX Location/Qualifiers

XX Key

XX Peptide

XX Peptide

XX Region

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site

PT and treatment of disorders associated with its activity including
 PT developmental, cell proliferative, immune, reproductive and
 PT cardiovascular disorders and infections -

PS Claim 1; Fig 2; 80pp; English.

XX This sequence represents human growth factor related molecule GFRP-2.
 CC cDNA encoding GFRP-1 was initially identified in a breast tissue
 CC cDNA library, and the present sequence is encoded by a consensus
 CC cDNA derived from several overlapping and/or extended cDNA clones.
 CC GFRP-2 is probably a CC chemokine and has chemical and structural
 CC homology with hTECK (20% identity). GFRP-2 and hTECK are both basic
 CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
 CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
 CC and nontumorous breast tissue. GFRP proteins (AAB03000-803003),
 CC nucleotides encoding them (AA52455-452458), GFRP agonists and
 CC antagonists may be used to treat a wide variety of diseases associated
 CC with increased or decreased expression or activity of GFRP proteins.
 CC Conditions which may be treated include developmental disorders, cell
 CC proliferative disorders (e.g., cancers), immune disorders (e.g.,
 CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
 CC disorders), cardiovascular disorders (e.g., arteriosclerosis) and
 CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
 CC proteins and nucleotides can be used in the diagnosis of such disorders.

XX Sequence 127 AA;

Query Match 100.0%; Score 97; DB 21; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KNGKGNVCHRRKHGK 16
 |||||
 DB 92 KNGKGNVCHRRKHGK 107

RESULT 9

AAB69175 standard; Protein: 127 AA.

XX AAB69175;

XX 26-APR-2001 (first entry)

XX Human G-protein coupled receptor GPR27 protein SEQ ID NO:2.

XX Human; G-protein coupled receptor; GPR27; antimicrobial; analgesic;
 KW cytosolic; antidiabetic; anorectic; antiaesthetic; antiparkinsonian;
 KW cardiatic; hypertensive; hypotensive; diuretic; osteopathic; antileptic;
 KW cerebroprotective; antileptogenic; antileptic; tranquilizer; nootropic;
 KW antidepressant; neuroleptic; anticonvulsant; vaccine; gene therapy;
 KW infection; pain; cancer; diabetes; obesity; anorexia; bulimia; asthma;
 KW Parkinson's disease; acute heart failure; hypotension; hypertension;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; vomiting;
 KW psychotic disorder; neurological disorder; dyslexia.

OS Homo sapiens.

XX WO200107482-A1.

XX 01-FEB-2001.

XX 21-JUL-2000; 2000WO-US19855.

XX 27-JUL-1999; 99US-0361564.

XX (SMK) SMITHKLINE BEECHAM CORP.

XX (SMK) SMITHKLINE BEECHAM PLC.

XX Testa TT, Vawter L;

XX WPI; 2001-191359/19.

XX Human growth factor related molecule protein useful for the diagnosis

```
DR N-PSDB; AAF59231.
XX
PT New G-protein coupled receptor GPR27 polypeptides and polynucleotides,
XX useful as vaccines or for treating diseases, e.g. infections, pain,
PT cancer, urinary retention, osteoporosis, stroke, psychotic and
PT neurological disorders -
XX
PS Claim 1; Page 29-30; 31pp; English.
XX
CC The present sequence represents a human G-protein coupled receptor,
XX designated GPR27. GPR27 has antimicrobial, analgesic, cytosolic,
CC antidiabetic, anorectic, antiasthmatic, antiparkinsonian, craditic,
CC hypertensive, hypotensive, diuretic, osteoprotic, cerebroprotective,
CC antitumor, antiallergic, antihemetic, tranquiliser, antidepressant,
CC neuroleptic, nootropic and anticonvulsant activities, and can be used
CC in producing vaccines and in gene therapy. The GPR27 protein and
CC polynucleotide are useful for treating certain diseases or as vaccines
CC against these diseases. These diseases include bacterial, fungal,
CC protozoan or viral (e.g. infections caused by HIV-1 or HIV-2) infections,
CC pain, cancers, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's
CC disease, acute heart failure, hypertension, urinary
CC retention, osteoporosis, angina pectoris, myocardial infarction, stroke,
CC ulcers, allergies, benign prostatic hyper trophy, migraine, vomiting,
CC psychosocial and neurological disorders (e.g. anxiety, schizophrenia, manic
CC depression, depression, delirium, dementia or mental retardation), or
CC dystonias (e.g. Huntington's disease or Gilles de la Tourette's
CC syndrome). The polynucleotide may also be used for chromosome
CC localisation studies, as tools for tissue expression studies, or as a
CC diagnostic reagent for detecting mutations in the associated gene.
XX
SQ Sequence 127 AA;
XX
Query Match 100.0%; Score 97; DB 22; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 1 KKGKGNVCHRRKKHCK 16
XXXXXXXXXXXXXXXXXXXX
DB 92 KKGKGNVCHRRKKHCK 107
XXXXXXXXXXXXXXXXXXXX
RESU.T 10
ABBO7728
ID ABBO7728 standard; Protein; 127 AA.
XX
XX ABB07728;
XX
XX 10-JUN-2002 (first entry)
DE Human mucosae-associated epithelial chemokine (Mec).
XX
XX Mucosae-associated epithelial chemokine; MEC; C-C chemokine receptor;
KW CG3; CCR10; anti-inflammatory; cytostatic; immunomodulator; anti-viral;
KW antibacterial; chemokine; human.
XX
XX Homo sapiens.
OS
XX WO200214532-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-US25734.
XX
XX 15-AUG-2000; 2000US-0638914.
XX
PA (MILL-) MILLENNIUM PHARM INC.
PA (STRD ) UNIV LEELAND STANFORD JUNIOR.
PI Butcher EC, Kunkel EJ, Pan J, Soler-Perran D;
DR WPI: 2002-269204/31.
DR N-PSDB; ABL40453.
XX
```

PT	Identifying modulators of mucosae-associated epithelial chemokine (MEC)
PT	receptors 3 or 10 (CCR3/10), useful for treating inflammatory diseases,
PT	comprises detecting formation of MEC-CCR3/10 complex or modulation of a
XX	MEC-induced response -
PS	
XX	
PS	Example 1; Fig 1A; 92pp; English.
CC	The invention relates to identifying agents that inhibit or promote the
CC	binding of a mammalian mucosae-associated epithelial chemokine (MEC) to
CC	a mammalian C-C chemokine receptor 3 (CCR3) or 10 (CCR10). The method
CC	involves: (a) detecting or measuring the formation of a complex between
CC	the MEC, and the CCR3 or CCR10; or (b) determining the ability of the
CC	test agent to inhibit or augment a MEC-induced response. An augmentation
CC	of complex formation, relative to a control, is indicative that the agent
CC	is a promoter. The method is useful for identifying modulators (e.g.
CC	inhibitors or promotor) of MEC-induced functions of CCR3 and/or CCR10.
CC	The inhibitors are useful for treating inflammatory diseases or
CC	conditions in a subject, e.g. oral inflammatory condition (e.g. Sjogren's
CC	sndrome or Behcet's syndrome), mastitis, chronic obstructive lung
CC	disease, asthma, inflammatory bowel disease (e.g. Crohn's disease,
CC	ulcerative colitis or celiac disease), Iga nephropathy or dermatitis
CC	herpetiformis. The promoters are useful for treating cancers (e.g. solid
CC	tumours or cutaneous T cell lymphoma), neoplastic disease, retinopathy,
CC	macular degeneration, bacterial infections, tuberculoid leprosy, viral
CC	infections, AIDS, neutropenias or bronchiectasis. The present sequence
CC	represents the human MEC protein.
SQ	
SQ	Sequence 127 AA:
Query Match	100.0%; Score 97; DB 23; Length 127;
Best Local Similarity	100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 KNGKGVNCHRRKHGK 16
Db	92 KNGKGVNCHRRKHGK 107
RESULT 11	
ABB08268	
ID	ABB08268 standard; Protein; 127 AA.
XX	
AC	ABB08268;
XX	
DT	20-MAY-2002 (first entry)
XX	
DE	Human mammary gland enriched chemokine.
XX	
KW	Human; MEC; mammary gland enriched chemokine; chemokine; tumour; cancer;
KM	cytostatic; antiinflammatory; inflammation.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..19
FT	/label=Leader_sequence
FT	20..127
FT	Protein /label=Mature_MEC
XX	
XX	US2002009735-A1.
PN	
XX	
PD	24-JAN-2002.
XX	
PF	21-MAR-2001; 2001US-0813492.
XX	
PR	23-MAR-2000; 2000US-191654P.
XX	
PA	(LABO/) LABOW M A.
PA	(MICK/) MICKANIN C S.
PA	(BHAT/) BHATTIA U.
XX	
PI	Labow MA, Mickanin CS, Bhatia U;
XX	

DR	WPI: 2002-187776/24.
DR	N-PSDB: ABA99025.
PT	Regulating tumour or adverse bodily reaction, involves providing
PT	therapeutic composition comprising a mammary gland chemokine, and
PT	providing the composition to the tumour or to the area of adverse
PT	reaction
PS	Claim 3; Fig 1; 11pp; English.
XX	
CC	The sequence represents human mammary gland enriched chemokine (MEC). The
CC	invention relates to a novel method for regulating a tumour or
CC	adverse bodily reaction, comprising providing a therapeutic composition
CC	having a mammary gland chemokine polypeptide. The polypeptide of the
CC	invention has cytosolic and antiinflammatory activity. The method of the
CC	invention is useful for regulating a tumour or adverse bodily reaction.
CC	The invention also provides a method useful for detecting a tumour using
CC	a probe comprising the polynucleotide or an antibody to the MEC. The
CC	adverse bodily reactions include cancer and inflammation.
XX	
SQ	Sequence 127 AA:
	Query Match 100.0%; Score 97; DB 23; Length 127;
	Best Local Similarity 100.0%; Pred. No. 1.8e-07;
	Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 KNGKGNVCHRRKHHGK 16
DB	92 KNGKGNVCHRRKHHGK 107
RESULT 12	
AAB47868	
ID	AAB47868 standard; Protein; 127 AA.
XX	
AC	AAB47868;
XX	
DT	15-APR-2002 (first entry)
XX	
DE	Full length placenta-derived C-C chemokine.
XX	
KW	Chemokine; placenta-derived; C-C chemokine; PLACC; inflammation;
KW	immune response; cell proliferation; asthma; malignancy; infection;
KW	adult respiratory distress syndrome; rheumatoid arthritis;
KW	lupus erythematosus; psoriasis; osteoarthritis; glomerulonephritis;
KW	osteoporosis; dermatomyositis; polymyositis; Addison's disease;
KW	Groves Disease; Crohn's disease; irritable bowel syndrome;
KW	atrophic gastritis; graft versus host disease; myasthenia gravis;
KW	multiple sclerosis; autoimmune thyroiditis; ulcerative colitis;
KW	atherosclerosis; pancreatitis.
KW	
KW	Homo sapiens.
OS	
XX	
EH	Key Location/Qualifiers
FT	Peptide 1..21
FT	/label= Signal_peptide
FT	22..127
FT	/label= C-C_chemokine
XX	
PN	WO200192301-A2.
XX	
PD	06-DEC-2001.
XX	
PD	23-MAY-2001; 2001WO-US16599.
XX	
PR	26-MAY-2000; 2000US-207578P.
XX	
XX	(INDV) UNIV INDIANA ADVANCED RES & TECHNOLOGY.
XX	
PA	Hromas RA;
PI	
XX	WPI: 2002-154522/20.
XX	N-PSDB: AA172313.
DR	

XX	New placenta-derived human C-C chemokine, useful for modulating an
PT	inflammatory or immune response, particularly for treating asthma,
PT	psoriasis, osteoarthritis, Graves Disease, Crohn's Disease or graft
PT	versus host disease
XX	
PS	Claim 1, Page 90-91; 97pp; English.
XX	
CC	This sequence shows a chemokine protein, which is a placenta-derived
CC	C-C chemokine (PLAC). The PLAC polypeptide or the chemokine-encoding
CC	polynucleotide, are useful for modulating an inflammatory/immune
CC	response, modulating the activity of PLAC and screening for modulators,
CC	and in stimulating cell proliferation. Particularly these are useful
CC	for treating asthma, adult respiratory distress syndrome, rheumatoid
CC	arthritis, lupus erythematosus, psoriasis, osteoarthritis,
CC	glomerulonephritis, osteoporosis, dermatomyositis, polymyositis,
CC	Addison's Disease, Graves Disease, Crohn's Disease, Irritable bowel
CC	syndrome, atrophic gastritis, graft versus host disease, myasthenia
CC	gravis, multiple sclerosis, autoimmune thyroiditis, ulcerative colitis,
CC	atherosclerosis or pancreatitis. Increasing the PLAC activity in the
CC	subject is useful where the subject may be exhibiting signs of a
CC	malnancy or infection or have an immunodeficiency.
XX	
SO	Sequence 127 AA;
Query Match	100.0%; Score 97; DB 23; Length 127;
Best Local Similarity	100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 KNGKGNVCHRRKKHNGK 16
Db	92 KNGKGNVCHRRKKHNGK 107
RESULT 13	
AAM25479	
ID	AAM25479 standard; Protein; 147 AA.
XX	
AC	AAM25479;
XX	
DT	16-OCT-2001 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:994.
XX	
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW	antinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW	antibacterial; endocrine; cardiac; central nervous system; virocid;
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
KW	antiagregant; haemostatic; vulnary; antitumor; osteopathic; eczema;
KW	dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW	neuroprotective; antidepressant; nocotropic; antiparkinsonian; infection;
KW	immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
KW	antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW	thrombocytopopenia; osteoporosis; severe combined immunodeficiency;
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX	neurological disorder.
OS	Homo sapiens.
XX	
PN	WO200153455-A2.
XX	
PD	26-JUL-2001.
XX	
PF	22-DEC-2000; 2000WO-US35017.
XX	
PR	23-DEC-1999; 99US-0471275.
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
XX	
PA	(HYSE-) HYSEQ INC.

Query	1	112	100.0%	Score 97	DB 22	Length 147			
Matches	16	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	1	112	100.0%	Score 97	DB 22	Length 147			
Db	112	112	100.0%	Score 97	DB 22	Length 147			
RESULT 14	ABP30895	ABP30895	standard; Protein; 1550 AA.						
XX	ABP30895;								
XX	02-JUL-2002	(first entry)							
XX	Streptococcus	polypeptide SEQ ID NO 10966.							
XX	Streptococcus	GAS; GBS; group B streptococcus; Streptococcus agalactiae;							
XX	group A streptococcus; Streptococcus pyogenes; antibacterial;								
XX	antiinflammatory; infection; vaccine; meningitis; gene therapy.								
XX	Streptococcus	agalactiae.							
XX	MO200234771-A2.								
XX	02-MAY-2002.								
XX	29-OCT-2001;	2001WO-GB04789.							
XX	27-OCT-2000;	2000GB-0026333.							
XX	24-NOV-2000;	2000GB-0028727.							
XX	07-MAR-2001;	2001GB-0005640.							
XX	(CHIR-) CHIRON SPA.								
XX	(GENO-) INST GENOMIC RES.								

PI	Telford J	Masignani V	Margarit Ros YI	Grandi G	Fraser C;
XX	Tettelin H;				
DR	WPI: 2002-352536/38.				
XX	N-PSDB: ABN71526.				
PT	New Streptococcus protein for the treatment or prevention of infection				
PT	or disease caused by Streptococcus bacteria, such as meningitis, and				
PT	for detecting a compound that binds to the protein -				
XX					
PS	Claim 1; Page 4196; 4525pp; English.				
XX					
CC	The invention relates to a protein (ABP25413-ABP30895) from group B				
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS				
CC	(Streptococcus pyogenes), comprising one of 5463 sequences (SI), given in				
CC	the specification. The proteins have antibacterial and antiinflammatory				
CC	activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and				
CC	antibodies that bind (1) are used in the manufacture of medicaments for				
CC	the treatment or prevention of infection or disease caused by				
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.				
CC	Nucleic acids encoding (1) are used to detect Streptococcus in a				
CC	biological sample. (1) is used to determine whether a compound binds to				
CC	(1). A composition comprising (1) or a nucleic acid encoding (1), may be				
CC	used as a vaccine or diagnostic composition. The disease caused by				
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic				
CC	acid encoding (1) may be used to recombinantly produce (1) and may be				
CC	used in gene therapy. Antibodies to (1) are used for affinity				
CC	chromatography, immunoassays, and distinguishing/identifying				
CC	Streptococcus proteins.				
XX					
SO	Sequence 1550 AA;				
	Query Match 50.0%; Score 48.5; DB 23; Length 1550;				
	Best Local Similarity 56.2%; Pred. No. 90;				
	Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1.				
OY	1 KNGKGNVCHRRKHGK 16				
	1:111:1:1111				
DB	1082 KDDKNGNI-KKOHGK 1096				
	RESULT 15				
	ABP29894				
ID	ABP29894 standard; Protein: 1570 AA.				
XX					
AC	ABP29894;				
XX					
DT	02-JUL-2002 (first entry)				
XX					
DE	Streptococcus polypeptide SEQ ID NO 8964.				
XX					
KM	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;				
KM	group A streptococcus; Streptococcus pyogenes; antibacterial;				
XX	antiinflammatory; Infection; vaccine; meningitis; gene therapy.				
OS	Streptococcus agalactiae.				
XX					
XX	MO200234771-A2.				
PN					
PD	02-MAY-2002.				
XX					
PF	29-OCT-2001; 2001MO-GB04789.				
XX					
PR	27-OCT-2000; 2000GB-0026333.				
XX	24-NOV-2000; 2000GB-0028727.				
XX	07-MAR-2001; 2001GB-0005640.				
PA	(CHIR-) CHIRON SPA.				
PA	(GENO-) INST GENOMIC RES.				
XX					
XX	Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;				
XX	Tettelin H;				

DR WPI; 2002-352536/38.
DR N-PSDB; ABN70525.

DR N-PSDB; ABN70525.

PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -

PT for detecting a compound that binds to the protein -

PT for detecting a compound that binds to the protein -

PS Claim 1; Page 4018; 4525pp; English.

XX

CC The invention relates to a protein (ABP25413-ABP30895) from group B

streptococcus/agalactiae) or group A streptococcus/GAS

CC (streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

activity. The proteins have antibacterial and antiinflammatory activity. The nucleic acids encoding (T) : ABN560AA-ABN71526 and

antibodies that bind (I) are used in the manufacture of medicaments for

the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyrr-*

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound

(I). A composition comprising (I) or a nucleic acid encoding (I), may be used for a method of detecting a target nucleic acid.

used as a vaccine or diagnostic composition. The disease caused by *Streptococcus* that is prevented or treated may be meningitis. Nucleic

acid encoding (I) may be used to recombinantly produce (I) and may be

used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

XX

SQ Sequence 1570 AA;

Best Local Similarity

Matches 9; Conserv

1 КНГКГНУСНРККНН

[illegible]

Db 1102 KDDKGNII-KKQHNHGK 1116

Search completed: January 14, 2003, 18:13:17
Job time : 12.2222 secs

Job time : 12.2222 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:07:25 ; Search time 9.1 Seconds

(without alignments)
407.566 Million cell updates/sec

Title: US-09-834-794A-3

Perfect score: 94

Sequence: 1 TEVSHHSIRRLERVMNC 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MNC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	88.3	135	11 Q91Y39	Q91Y39 ratius norv
2	54	57.4	1591	3 Q9HFW1	Q9HFW1 ashbya goss
3	48	51.1	1245	4 Q9NPK3	Q9NPK3 homo sapien
4	48	51.1	1246	4 Q76046	Q76046 homo sapien
5	48	51.1	1246	4 Q96BC0	Q96BC0 homo sapien
6	46	48.9	161	10 Q8RY03	Q8RY03 arabidopsis
7	46	48.9	421	10 Q8SJ32	Q8SJ32 arabidopsis
8	46	48.9	798	2 Q50255	Q50255 rhizobium e
9	45	47.9	292	5 Q8T9V0	Q8T9V0 aedes aegypt
10	45	47.9	318	2 Q9L7Y6	Q9L7Y6 pseudomonas
11	45	47.9	342	16 Q985S7	Q985S7 rhizobium 1
12	44	46.8	232	10 Q43707	Q43707 zea mays (m
13	44	46.8	466	10 Q9M711	Q9M711 plasm sativ
14	44	46.8	818	17 Q8TYU9	Q8TYU9 methanopyru
15	43	45.7	268	16 Q9RUM5	Q9RUM5 delnoccocus
16	43	45.7	335	8 Q9GEZ1	Q9GEZ1 ipomopsis a

17	43	45.7	351	16 Q8X2S6	Q8X2S6 ralsstonia s
18	43	45.7	581	10 Q9L1B8	Q9L1B8 arabidopsis
19	43	45.7	584	10 Q43866	Q43866 arabidopsis
20	43	45.7	1028	10 Q9FGH7	Q9FGH7 arabidopsis
21	43	45.7	1059	16 Q9KFM4	Q9KFM4 bacillus ha
22	42.5	45.2	430	11 Q991V8	Q991V8 mus musculu
23	42.5	45.2	494	11 Q08612	Q08612 mus musculu
24	42.5	45.2	1090	11 Q922B6	Q922B6 mus musculu
25	42.5	45.2	1147	11 Q9OWY8	Q9OWY8 mus musculu
26	42	44.7	105	11 Q9D018	Q9D018 mus musculu
27	42	44.7	270	16 Q8XEM4	Q8XEM4 salmonella
28	42	44.7	271	16 Q8XA12	Q8XA12 escherichia
29	42	44.7	388	5 Q9VMB3	Q9VMB3 drosophila
30	42	44.7	715	5 Q9N8R8	Q9N8R8 trypanosoma
31	42	44.7	1108	5 Q9VHD8	Q9VHD8 drosophila
32	42	44.7	1108	5 Q9TYG5	Q9TYG5 drosophila
33	41	43.6	182	5 Q9VJ99	Q9VJ99 drosophila
34	41	43.6	293	16 Q8ZAA7	Q8ZAA7 yersinia pe
35	41	43.6	367	10 Q8VZK0	Q8VZK0 arabidopsis
36	41	43.6	457	17 Q8ZT57	Q8ZT57 pyrobaculum
37	41	43.6	501	16 Q92R57	Q92R57 rhizobium m
38	41	43.6	508	10 Q9FTX1	Q9FTX1 arabidopsis
39	41	43.6	548	16 Q9PK16	Q9PK16 chlamydia m
40	41	43.6	732	12 Q9DUC7	Q9DUC7 te virus. o
41	41	43.6	795	11 Q8R3X0	Q8R3X0 mus musculu
42	41	43.6	1236	11 Q70349	Q70349 mus musculu
43	41	43.6	1313	4 Q9HCK6	Q9HCK6 homo sapien
44	41	43.6	1743	4 Q9BYP7	Q9BYP7 homo sapien
45	41	43.6	1800	4 Q8TCK6	Q8TCK6 homo sapien

ALIGNMENTS

RESULT 1					
Q91Y39		PRELIMINARY:	PRT:	135 AA.	
AC Q91Y39					
DT 01-DEC-2001 (TREMBLrel. 19, Created)					
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)					
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)					
DE CC Chemokine CCL28.					
OS Rattus norvegicus (Rat).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OX NCBI_TaxID=10116;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=Salivary gland;					
RA Beuken E., Gruijthuisen Y.K., Brugeman C.A., Vink C.;					
RT "Rattus norvegicus CC chemokine CCL28 (rCCL28) mRNA.";					
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.					
DR EMBL: AF361490; AAK52773.1; -					
DR InterPro: IPR001811; Chemokine_IL8.					
SQ SEQUENCE 135 AA; 15096 MW; AFR79E0EF7EA8A64 CRC64;					
Query Match	88.3%;	Score 83;	DB 11;	Length 135;	
Best Local Similarity	88.9%;	Pred. No. 3.4e-06;			
Matches	16;	Conservative	0;	Mismatches	2;
				Indels	0;
				Gaps	0;
QY 1 TEVSHHSIRRLERVMNC 18					
DB 32 TEVSHHSIRRLERVMNC 49					
RESULT 2					
Q9HFW1		PRELIMINARY:	PRT:	1591 AA.	
ID Q9HFW1					
AC Q9HFW1					
DT 01-MAR-2001 (TREMBLrel. 16, Created)					
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
DE YN1297C.					

GN VNL297C.
 OS Ashbya gossypii (Eremothecium gossypii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Eremotheciaceae; Eremothecium.
 OX NCBI_TaxID=33169;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20534876; PubMed=11082049;
 RA Ayad-Durieux Y., Knechtle P., Golf S., Dietrich F., Philippsen P.;
 RT "A PAK-like protein kinase is required for maturation of young hyphae
 and septation in the filamentous ascomycete Ashbya gossypii.";
 RL J. Cell Sci. 113:4563-4575(2000).
 SO EMBL: AF286114; AAC17721.1; -;
 SQ SEQUENCE 1591 AA; 181270 MW; 3E9A063EFC40C2DF3 CRC64;

Query Match 57.4%; Score 54; DB 3; Length 1591;
 Best Local Similarity 62.5%; Pred. No. 2.9;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 VSHHSIRRLERVMC 18
 ||:|:||||:|
 Db 1268 VSYHASHMILLERINAC 1283

RESULT 3
 O9NPK3 PRELIMINARY; PRT; 1245 AA.
 AC O9NPK3;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE DJ347.7 (superkiller virulicidic activity 2 (S. cerevisiae
 DE homo109)-like (SKI2W)).
 GN SKI2L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Barlow K.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL049547; CAB89307.1; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1245 AA; 137623 MW; 981DDF3DADCFEFD CRC64;

Query Match 51.1%; Score 48; DB 4; Length 1245;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TEVSHHSIRRLERVN 16
 ||:|:||||:|
 Db 818 TETQHMIOIRIMESVN 833

RESULT 4
 O76046 PRELIMINARY; PRT; 1246 AA.
 AC O76046;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Putative RNA helicase SKI2W.
 GN SKI2W.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95334363; PubMed=7610041;
 RA Dangel A.W., Shen L., Mendoza A.R., Wu L.C., Yu C.Y.;
 RT "Human helicase gene SKI2W in the HLA class III region exhibits
 striking structural similarities to the yeast antiviral gene SKI2 and
 to the human gene KIA0052: emergence of a new gene family.";
 RL Nucleic Acids Res. 23:2120-2126(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99017970; PubMed=9799600;
 RA Yang Z., Shen L., Dangel A.W., Wu L.C., Yu C.Y.;
 RT "Four ubiquitously expressed genes, RD (D6S45)-SKI2W (SKI2L)-DOM3Z-
 RP1 (D6S60E), are present between complement component genes factor B
 RT and C4 in the class III region of the HLA.";
 RL Genomics 53:338-347(1998).
 DR EMBL: AF059675; AAC78607.1; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1246 AA; 137794 MW; FCDE30D09D6B0F8 CRC64;

Query Match 51.1%; Score 48; DB 4; Length 1246;
 Best Local Similarity 56.2%; Pred. No. 23;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TEVSHHSIRRLERVN 16
 ||:|:||||:|
 Db 819 TETQHMIOIRIMESVN 834

RESULT 5
 O96BC0 PRELIMINARY; PRT; 1246 AA.
 AC O96BC0;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Superkiller virulicidic activity 2 (S. cerevisiae homolog)-like.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC015758; AAH15758.1; -;
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00270; DEAD; 1.
 DR Pfam: PF00271; helicase_C; 1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 1246 AA; 137793 MW; 6FF44975E305B850 CRC64;

Query Match 51.1%; Score 48; DB 4; Length 1246;
 Best Local Similarity 56.2%; Pred. No. 23;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TEVSHHSIRRLERVN 16
 ||:|:||||:|
 Db 819 TETQHMIOIRIMESVN 834

RESULT 6
 O8RY03 PRELIMINARY; PRT; 161 AA.
 ID O8RY03

```

AC 08RY03;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 18.6 kDa protein.
GN AT2G05300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shim P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis full length cDNA clones."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY080586; AAL85963.1; -
KM Hypothetical protein. 161 AA; 18599 MW; 9C0LADICCFADP68D CRC64;
SQ SEQUENCE 161 AA; 18599 MW; 9C0LADICCFADP68D CRC64;

Query Match 48.9%; Score 46; DB 10; Length 161;
Best Local Similarity 57.1%; Pred. No. 6.2;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 SHHSRRLRLRYNM 17
DB 73 SHLSRKAMERYSM 86
II:::II:::

RESULT 7
OY 09SJ32 PRELIMINARY; PRT; 421 AA.
ID 09SJ32;
AC 09SJ32;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE At3g05300 protein.
GN AT2G05300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC MEDLINE=20083487; PubMed=10617197;
RX Lin X., Kaul S., Rounsley S.D., Shea T.P., Beutlo M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.D., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Sher M., Vanaken S.E., Umeyam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768(1999).
RN [2];
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC007018; AAD29066.1; -
DR InterPro: IPR001810; F-box.
DR Pfam: PF00646; F-box; 1.
DR SMART: SM00256; FBOX; 1.
DR PROSITE: PS0181; FBOX; 1.

```

```

SO SEQUENCE 421 AA; 48188 MW; A77E3CB15A1A1 CRC64;

Query Match 48.9%; Score 46; DB 10; Length 421;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 4 SHHSRRLRLRYNM 17
DB 333 SHLSRKAMERYSM 346
II:::II:::

RESULT 8
OY 050255 PRELIMINARY; PRT; 798 AA.
ID 050255;
AC 050255;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Stomatin like protein (slp).
OS Rhizobium etli.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=29449;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=TAL182;
RX MEDLINE=98455831; PubMed=9782511;
RA You Z., Gao X., Ho M.M., Borthakur D.;
RT "A stomatin-like protein encoded by the slp gene of Rhizobium etli is
required for nodulation competitiveness on the common bean."
RL Microbiology 144:2619-2627(1998).
DR EMBL: AF034831; AAC64871.1; -
DR InterPro: IPR001633; EAF.
DR InterPro: IPR003018; GAF.
DR InterPro: IPR00160; GDEF.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR00700; PAS-assoc-C.
DR InterPro: IPR00014; PAS_domain.
DR Pfam: PF00563; EAF; 1.
DR Pfam: PF01590; GAF; 1.
DR Pfam: PF00990; GDEF; 1.
DR Pfam: PF00785; PAC; 1.
DR SMART: SM00267; DUF1; 1.
DR SMART: SM00552; DUF2; 1.
DR SMART: SM00065; GAF; 1.
DR SMART: SM00086; PAC; 1.
DR TIGRFAMs: TIGR00254; GDEF; 1.
DR TIGRFAMs: TIGR00229; sensory_box; 1.
SO SEQUENCE 798 AA; 89747 MW; 7644177MAA7727A1 CRC64;

Query Match 48.9%; Score 46; DB 2; Length 798;
Best Local Similarity 61.5%; Pred. No. 31;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 VSHHSRRLRLRY 15
DB 644 VTHHTTTRLLEQV 656
II:::II:::

RESULT 9
OY 08T9V0 PRELIMINARY; PRT; 292 AA.
ID 08T9V0;
AC 08T9V0;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative secreted protein.
OS Aedes aegypti (Yellowfever mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
OC Aedes.
OX NCBI_TaxID=7159;
RN [1];

```

RP SEQUENCE FROM N.A.
RC STRAIN-BLACK EYE; TISSUE-SALIVARY GLAND;
RA Valenzuela J.G., Pham V.M., Garfield M.K., Francischetti I.M.,
RA Ribeiro J.M.C.;
RT "Toward the silome of the adult female mosquito *Aedes aegypti*."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF46596; AAL76019.1; -
SQ SEQUENCE 292 AA; 33720 MW; 47DC62EF2A3F0C25 CRC64;

Query Match 47.9%; Score 45; DB 5; Length 292;
Best Local Similarity 38.9%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLERVMNC 18
DB 67 TQIKRIKHLEKCNLC 84
:::|||||::|

RESULT 10
O9L7Y6 PRELIMINARY; PRT; 318 AA.
AC O9L7Y6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE BENR.
GN BENR.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=503;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20507802; PubMed-11053377;
RA Cowles C.E., Nichols N.N., Harwood C.S.;
RT "Benr, a xyls homologue, regulates three different pathways of
RT aromatic acid degradation in *Pseudomonas putida*."
RL J. Bacteriol. 182:6338-6346(2000).
CC -1 SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: AF218267; AAF63447.1; -
DR InterPro: IPR000005; HTHARAC.
DR Pfam: PF00165; HTH_ARAC; 2.
DR PRINTS: PRO0032; HTHARAC.
DR SMART: SM00342; HTH_ARAC; 1.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
KW DNA-binding; Transcription regulation.
SQ SEQUENCE 318 AA; 36451 MW; 46E4162941C2A02 CRC64;

Query Match 47.9%; Score 45; DB 2; Length 318;
Best Local Similarity 44.4%; Pred. No. 18;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLERVMNC 18
DB 255 TTPKHVYRQKRLERVHAC 272
|:::|||||::|

RESULT 11
O985S7 PRELIMINARY; PRT; 342 AA.
AC O985S7;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Nucleotide sugar epimerase.
GN MUR7549.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-MAFE30309;
RX MEDLINE-21082930; PubMed-11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003012; BAB53985.1; -
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase_1.
DR PRINTS: PR01497; SHALCHANNEL.
KW Complete proteome.
SQ SEQUENCE 342 AA; 37527 MW; BCFB9652FA161556 CRC64;

Query Match 47.9%; Score 45; DB 16; Length 342;
Best Local Similarity 58.3%; Pred. No. 19;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 VSHHSRRLER 14
DB 12 IGYHVARRLER 23
:::|||||::|

RESULT 12
O43707 PRELIMINARY; PRT; 232 AA.
ID O43707;
AC O43707;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE (clone wus11032) mRNA sequence.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. BERKELEY FAST; TISSUE-COLEOPTILE;
RX MEDLINE-94211899; PubMed-7909162;
RA Pesche V.M., Sachs M.M.;
RT "Characterization and expression of transcripts induced by oxygen
RT deprivation in maize (*Zea mays* L.)."
RL Plant Physiol. 104:387-394(1994).
DR EMBL: L14835; AAA20404.1; -
SQ SEQUENCE 232 AA; 23995 MW; 712D0E20B1DB1F CRC64;

Query Match 46.8%; Score 44; DB 10; Length 232;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 VSHHSRRLERVMNC 17
DB 104 LAHHLPRRLERLRY 118
:::|||||::|

RESULT 13
O9M711 PRELIMINARY; PRT; 466 AA.
ID O9M711;
AC O9M711;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative Hsipro-1 homolog.
GN HSI1PRO-1.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

CC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI.
 RA Kang J.-G., Park C.-M.;
 RT "Isolation of a cDNA clone encoding a putative Hsipro-1 homolog from
 pea."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF148506; AAF67003.1; -
 DR ANU-2PAGE: 09M711; -
 SO SEQUENCE 466 AA; 53052 MW; 6B8B7AB44DC0927 CRC64;
 Query Match 46.8%; Score 44; DB 10; Length 466;
 Best Local Similarity 63.6%; Pred. No. 39;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 OY 6 HISRLRLRVN 16
 DB 254 HSKRLRLRIS 264
 RESULT 14
 O8RYU9 PRELIMINARY; PRT; 818 AA.
 ID O8RYU9
 AC O8RYU9;
 DT 01-JUN-2002 (TREMBLER, 21, Created)
 DT 01-JUN-2002 (TREMBLER, 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLER, 21, Last annotation update)
 DE Mg-chelatase subunit ChlI and ChlD (Moxr-like ATPase and vnf
 DE domain).
 GN CHL1/CHLD OR MK0192.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 CC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE-21927647; PubMed-11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natsale D.A., Kozlov I.B., Tatusev R.L., Wolf Y.I., Stetter K.O.,
 RA Malysk A.G., Kozlov E.V., Kozlov S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens."
 RL PROC. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AE010318; AA001409.1; -
 KW Complete proteome.
 SO SEQUENCE 818 AA; 91451 MW; 8PE658F4E241B7D2 CRC64;
 Query Match 46.8%; Score 44; DB 17; Length 818;
 Best Local Similarity 57.1%; Pred. No. 69;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 TEVSHHISRLRLR 14
 DB 570 TEGFHNAKRLSER 583
 RESULT 15
 O9RUM5 PRELIMINARY; PRT; 268 AA.
 ID O9RUM5
 AC O9RUM5;
 DT 01-MAY-2000 (TREMBLER, 13, Created)
 DT 01-MAY-2000 (TREMBLER, 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)
 DE Hypothetical protein DR1265.
 GN DR1265.
 OS Deinococcus radiodurans.
 CC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococaceae; Deinococcus.
 OX NCBI_TaxID=1299;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI.
 RX MEDLINE-20036896; PubMed-10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI."
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001974; AAF10841.1; -
 DR TIGR: DR1265; -
 DR InterPro: IPR000182; GCN5acetyltransf.
 DR Pfam: PF00583; Acetyltransf; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 268 AA; 29493 MW; B262702F59A106B2 CRC64;
 Query Match 45.7%; Score 43; DB 16; Length 268;
 Best Local Similarity 57.1%; Pred. No. 33;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 1 TEVSHHISRLRLR 14
 DB 240 TEASHNARRIFER 253

Search completed: January 14, 2003, 18:15:28
 Job time : 11.1 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:08:30 ; Search time 4.4 Seconds
(without alignments)
393.277 Million cell updates/sec

Title: US-09-834-794A-3

Perfect score: 94

Sequence: 1 TEVSHHISRRLERVMNC 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	52.1	400	1 ZBBER4	44.1K zinc-binding
2	48	51.1	1246	2 S56752	helicase SKI2W - h
3	46	48.9	421	2 H84466	hypothetical prote
4	44	46.8	232	2 T04354	hypothetical prote
5	43	45.7	268	2 C75416	hypothetical prote
6	43	45.7	584	2 S37212	beta-fructofuranos
7	43	45.7	1059	2 G83706	lambdoidic mersac
8	42.5	45.2	1147	2 T42627	ADP-ribosylation f
9	42	44.7	270	2 AR0514	DnaJ-like protein
10	42	44.7	271	2 G64726	yadh protein - Esc
11	42	44.7	271	2 D50636	probable DNA bindi
12	42	44.7	271	2 D85487	probable DNA bindi
13	42	44.7	1108	2 JC4037	alpha-mannosidase
14	41	43.6	293	2 AFO475	LYR-family transc
15	41	43.6	548	2 B81698	pyrophosphate-fruc
16	41	43.6	700	2 T41401	spore outgrowth fa
17	41	43.6	1124	2 S06836	phlyochrome - gard
18	40.5	43.1	271	2 A25669	PdPf-related trans
19	40.5	43.1	558	1 MNXRR	RNA 10 protein - r
20	40	42.6	158	2 E96733	hypothetical prote
21	40	42.6	222	2 S18246	xyli protein - pse
22	40	42.6	253	2 H72017	snlycerol-3-p acy
23	40	42.6	253	2 G86605	snlycerol-3-p acy
24	40	42.6	269	2 T22778	hypothetical prote
25	40	42.6	295	2 F72082	had homolog hydrol
26	40	42.6	295	2 E86541	HAD type hydrolase
27	40	42.6	299	2 T19564	hypothetical prote
28	40	42.6	299	2 T35407	hypothetical prote
29	40	42.6	429	2 JC4636	rRNA endonuclease

30	40	42.6	436	2 A69223	formylmethanofuran
31	40	42.6	920	2 T41282	probable transcrip
32	40	42.6	1129	2 D84022	ATP-dependent nucl
33	39.5	42.0	226	1 TVMWS	PGF-related trans
34	39.5	42.0	241	1 PRHUG2	platelet-derived g
35	39	41.5	66	2 T18027	hypothetical prote
36	39	41.5	201	2 E72304	amidotransferase -
37	39	41.5	207	2 E69085	conserved hypothet
38	39	41.5	229	2 T33627	hypothetical prote
39	39	41.5	349	2 A97087	membrane associate
40	39	41.5	366	1 B64455	probable phosphos
41	39	41.5	432	2 T20929	hypothetical prote
42	39	41.5	448	2 E89936	hypothetical prote
43	39	41.5	510	2 A10490	sugar transport AT
44	39	41.5	667	2 T26136	hypothetical prote
45	39	41.5	1757	2 T05204	hypothetical prote

ALIGNMENTS

RESULT 1

ZBBER4 44.1K zinc-binding protein - Ictalurid herpesvirus 1 (strain auburn 1)

C:Species: Ictalurid herpesvirus 1

A:Note: host Ictalurus punctatus (channel catfish)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999

C:Accession: D36794

R:Davidson, A.J.

Submitted to GenBank, January 1992

A:Description: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A36804

A:Accession: D36794

A:Molecule type: DNA

A:Residues: 1-400 <DAY>

A:Cross-references: GB:M75136; NID:q331209; PIDN:AAA8180.1; PID:q331286

R:Davidson, A.J.

Virology 186, 9-14, 1992

A:Title: Channel catfish virus: a new type of herpesvirus.

A:Reference number: A39447; MUID:92087490; PMID:1727613

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 78

C:Superfamily: Ictalurid herpesvirus 44.1K zinc binding protein

C:Keywords: zinc finger

Query Match 52.1%; Score 49; DB 1; Length 400;

Best Local Similarity 52.9%; Pred. No. 2.3;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 EVSHHISRRLERVMNC 18

DB 190 KVQRHISRRLERVMNC 206

RESULT 2

S56752

helicase SKI2W - human

N:Alternate names: I70A; nucleolar protein SKI2 homolog; SKI2W; SKI2V; superkiller v1

C:Species: Homo sapiens (man)

C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 02-Mar-2001

C:Accession: S56752; A56003

R:Dangel, A.W.; Shen, L.; Mendoza, A.R.; Wu, L.C.; Yu, C.Y.

Nucleic Acids Res. 23, 2120-2126, 1995

A:Title: Human helicase gene SKI2W in the HLA class III region exhibits striking stru

Y.

A:Reference number: S56752; MUID:95334363; PMID:7610041

A:Accession: S56752

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-1246 <DNM>

A:Cross-references: EMBL:Z48796; NID:9914852; PIDN:CAA88733.1; PID:94379045

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1995
R:Lee, S.-G.; Lee, I.; Park, S.H.; Kang, C.; Song, K.
Genomics 25, 660-666, 1995
B:Title: Identification and characterization of a human cDNA homologous to yeast SKT2.
A:Reference number: A56003; MUID:95278935; PMID:7759100
A:Accession: A56003
A:Molecule type: mRNA
A:Residues: 'GGGTAARRSOR', 240-365, 'T', 367-561, 'PHVPSCPMWCSPSPGAAVST', 564-622, 'H', 624-913
A:Cross-references: GB:U09877; NID:G154826
C:Genetics:
A:Gene: GDB:SKIV2L; SKI2W; 170A; SKI2
A:Cross-references: GDB:542886; OMIM:600478
A:Map position: 6p21-6p21
C:Superfamily: yeast probable SKT2 protein YTL050W
C:Keywords: ATP; nucleotide binding; P-loop
E:413-424/Region: nucleotide-binding motif A (P-loop)
E:423-426/Region: nucleotide-binding motif B

```

Query Match      51.1%;  Score 48;  DB 2;  Length 1246;
Best Local Similarity 56.2%;
Matches 9;  Conservative 2;  Mismatches 5;  Indels 0;  Gaps 0.
OY 1 TEVSHHISRRLRYN 16
   ||| | | | | | | |
Db 819 TETOMIORIMESYN 834

```

RESULT 3
 H84466
 hypothetical protein At2g05300 [imported] - Arabidopsis thaliana
 C.Species: Arabidopsis thaliana (mouse-ear cress)
 C.Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C.Accession: H84466
 R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffett, K.S.; Cronan, L.A.; Shen, M.; Vankken, S.E.; Mayama, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A.Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A.Reference number: A84420; MUID:20083487; PMID:10617197
 A.Accession: H84466
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-421 <SMD>
 A.Cross-references: GB:AE002093; NID:g4755199; PIDN:AD29066.1; GSPDB:GN00139
 C.Genetics:
 A.Gene: At2g05300
 A.Map position: 2

Query Match	48.9%	Score 46	DB 2	Length 422
Best Local Similarity	57.1%	Pred. No. 7	8	
Matches	8	Conservative	4	Mismatches 2
				Indels 0
				Gaps 0
QY	4	SHHISRRLLERYNM	17	
		: : :		
DB	333	SHQLSRKAMERYSM	346	

```

RESULT 4
T04354
hypothetical protein - maize
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04354
R:Paschke, V.M.; Sachs, M.M.
Plant Physiol. 104, 387-394, 1994
A:Title: Characterization and expression of transcripts induced by oxygen deprivation in
A:Reference number: Z15305; MUID:94211899; PMID:7909162
A:Accession: T04354
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1*232 <PES>
A:Cross-references: EMBL:L14835; NID:g2993911; PIDN:AAA20404.1; PID:g520361

```

A: Experimental source: cultivar Berkeley Fast; coleoptile
A: Note: Induced by oxygen deprivation

Query Match	46.8%;	Score 44;	DB 2;	Length 233;
Best Local Similarity	46.7%;	Pred. No. 9;		
Matches	7;	Conservative	5;	Mismatches 3; Indels 0; Gaps 0;

```
QY      3 VSHHSRLLERVNM 17
        ::||: ||||: :
Db     104 LAHHLPHRLLERLV 118
```

RESULT 5
 C75416
 hypothetical protein - *Delnoccocus radiodurans* (strain R1)
 C1Species: *Delnoccocus radiodurans*
 C1Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
 C1Accession: C75416
 R1White, O. J., Eelsen, J. A.; Heldelberg, J. F.; Hickey, E. K.; Peterson, J. D.; Dodson, R. J.;
 M. M. Shen, M. J.; Vamathevan, J. J.; Lam, P. J.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S. J. Smith, H. O.; Venter, J. C.; Fraser, C. M.
 Science 286, 1571-1577, 1999
 A1Title: Genome sequence of the radioresistant bacterium *Delnoccocus radiodurans* R1.
 A1Reference number: A75250; MUID:20036896; PMID:10567266

A:Molecule type: DNA
A:Residues: 1-268 <WHF>
A:Cross-references: GB:AE001974; GB:AE000513; NID:56459010; PIDN:AAF10841.1; PID:5645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1265
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1265

Query Match	45.7%	Score 43	DB 2	Length 266
Best Local Similarity	57.1%	Pred. No. 15		
Matches	8	Conservative	2	Mismatches 4
				Indels 0
				Gaps 0
QY	1	TEVSHHISRLRLR	14	
		: : :		
DB	240	TEASHNHAMRIER	253	

RESULT 6
S37212
beta-fructofuranosidase (EC 3.2.1.26) 1, 66.2K - Arabidopsis thaliana
M.Alternate names: Invertase
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 22-Jun-1999
C.Accession: S37212
R.Schweibel-Bugne, N.; El Mtili, N.; Krivitzky, M.; Jean-Jacques, I.; Williams, J.H.H.
submitted to the EMBL data library September 1993
A.Description: Molecular cloning of a gene and cognate cDNA encoding a cell wall invertase
A.Reference number: S37212
A.Accession: S37212
A.Molecule type: mRNA
A.Residues: 1-584 <SCCH>
A.Cross-references: EMBL:X74514; NTD:g402739; PIDD:CA52619.1; PID:g402740
C.Superfamily: beta-fructofuranosidase
C.Keywords: glycosidase; hydrolase

Query Match	45.7%	Score 43	DB 2	Length 584
Best Local Similarity	46.7%	Pred. No. 35		
Matches 7	Conservative 3	Mismatches 5	Indels 0	Gaps 0
QY	2	EVSHHSRRRLERN	16	
		: :		
Db	27	EASHSVYKRLQSTN	41	

RESULT 7
G83706

Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLERVM 17
 |||:::|:|:|:
 Db 73 TEADHIASQLMDRML 89

RESULT 12

D85487

Probable DNA binding protein yabH [Imported] - Escherichia coli (strain O157:H7, subtype
 C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: D85487

R:Perne, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 11ter, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousts, K.; Apodaca,
 Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: D85487

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-271 <STO>

A:Cross-references: GB:AE005174; NID:g12512750; PIDN:AAG54360.1; GSPDB:GN00145; UMG:200

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

C:Superfamily: dnaJ amino-terminal homology

Query Match 44.7%; Score 42; DB 2; Length 271;
 Best Local Similarity 41.2%; Pred. No. 23;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLERVM 17
 |||:::|:|:|:
 Db 73 TEADHIASQLMDRML 89

RESULT 13

JC4037

alpha-mannosidase II, Golgi - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Jan-1995 #sequence_revision 21-Jun-2002 #text_change 21-Jun-2002

C:Accession: JC4037, S47019

R:Forster, J.M.; Yuckin, B.; Lockyer, A.E.; Roberts, D.B.

A:Title: Cloning and sequence analysis of gmit, a Drosophila melanogaster homologue of

A:Reference number: JC4037; MUID:95196999; PMID:7890162

A:Accession: JC4037

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1108 <POS>

A:Cross-references: EMBL:X77652; NID:g517480; PIDN:CAA54732.1; PID:g517481

A:Experimental source: embryo

C:Genetics:

A:Gene: FlyBase:alpha-Man-II

A:Cross-references: FlyBase:FBgn0011740

C:Keywords: Transmembrane protein

F:8-27/Domain: transmembrane #status predicted <TM>

Query Match 44.7%; Score 42; DB 2; Length 1108;
 Best Local Similarity 61.5%; Pred. No. 1e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 HISRLRLERVMC 18
 |||:::|:|:|:
 Db 953 HIRYRLERVMNC 965

RESULT 14

AF0475

Yersin-family transcription regulatory protein YPO3904 [Imported] - Yersinia pestis (strai
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AF0475

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davis, R.M.; Davis, P.; Dougan, G

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barr

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AF0475

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-293 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC93370.1; PID:g15981816; GSPDB:GN00175

C:Genetics:

A:Gene: YPO3904

Query Match 43.6%; Score 41; DB 2; Length 293;
 Best Local Similarity 44.4%; Pred. No. 36;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 TEVSHHSRRLERVMC 18
 |||:::|:|:|:
 Db 217 TTTSAHLTRQLLEWTGSC 234

RESULT 15

E81698

pyrophosphate-fructose 6-phosphate 1-phosphotransferase, beta chain TC0479 [Imported]

C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000

C:Accession: E81698

R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heldelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: E81698

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-548 <RET>

A:Cross-references: GB:AE002316; GB:AE002160; NID:g7190515; PIDN:AAF39325.1; PID:g719

A:Experimental source: strain N19g (Mopn)

C:Genetics:

A:Gene: TC0479

Query Match 43.6%; Score 41; DB 2; Length 548;
 Best Local Similarity 58.3%; Pred. No. 70;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 VSHHSRRLER 14
 |||:::|:|:|:
 Db 287 LSHHIAHLIRK 298

Search completed: January 14, 2003, 18:16:20
 Job time: 6.4 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 18:11:25 ; Search time 4 Seconds
(without alignments)
132.403 Million cell updates/sec

Title: US-09-834-794A-3
Perfect score: 94
Sequence: 1 TEVSHHSIRRLRYVMNC 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	18	4	US-09-146-580-3
2	94	100.0	104	4	US-09-146-580-2
3	94	100.0	127	4	US-09-146-580-1
4	40	42.6	293	2	US-09-049-671-1
5	40	42.6	293	3	US-09-295-068-1
6	40	42.6	302	3	US-09-221-235-2
7	40	42.6	302	3	US-09-221-928-2
8	40	42.6	302	4	US-09-221-527-2
9	40	42.6	302	4	US-09-221-236-2
10	40	42.6	302	4	US-09-221-416-2
11	40	42.6	302	4	US-09-221-245-2
12	40	42.6	302	4	US-09-163-115-2
13	40	42.6	302	4	US-09-221-528-2
14	40	42.6	302	4	US-09-593-553-2
15	40	42.6	302	4	US-09-221-237-2
16	39.5	42.0	102	2	US-08-804-953-2
17	39.5	42.0	109	1	US-08-094-079-2
18	39.5	42.0	109	2	US-08-804-953-3
19	39.5	42.0	109	3	US-08-691-794-4
20	39.5	42.0	109	5	PCT-US91-02612-1
21	39.5	42.0	109	5	PCT-US93-02612-1
22	39.5	42.0	109	6	US-08-257-494D-1
23	39.5	42.0	119	2	US-08-804-953-2
24	39.5	42.0	120	6	5428135-2
25	39.5	42.0	129	6	5428135-6
26	39.5	42.0	146	3	US-08-989-251-2
27	39.5	42.0	146	3	US-08-989-251-2

28	39.5	42.0	146	3	US-09-340-250-2	Sequence 2, Appl1
29	39.5	42.0	146	3	US-09-340-250-25	Sequence 25, Appl1
30	39.5	42.0	146	3	US-09-528-108-2	Sequence 2, Appl1
31	39.5	42.0	146	4	US-09-528-108-25	Sequence 25, Appl1
32	39.5	42.0	160	1	US-08-094-079-1	Sequence 1, Appl1
33	39.5	42.0	190	3	US-08-867-352-25	Sequence 25, Appl1
34	39.5	42.0	205	3	US-08-989-251-27	Sequence 27, Appl1
35	39.5	42.0	205	3	US-08-989-251-37	Sequence 37, Appl1
36	39.5	42.0	205	3	US-09-340-250-27	Sequence 27, Appl1
37	39.5	42.0	205	3	US-09-340-250-37	Sequence 37, Appl1
38	39.5	42.0	205	4	US-09-528-108-27	Sequence 27, Appl1
39	39.5	42.0	205	4	US-09-528-108-37	Sequence 37, Appl1
40	39.5	42.0	220	6	5175255-4	Patent No. 5175255
41	39.5	42.0	226	6	5498600-2	Patent No. 5498600
42	39.5	42.0	241	1	US-08-387-845-4	Sequence 4, Appl1
43	39.5	42.0	241	2	US-08-999-811-6	Sequence 6, Appl1
44	39.5	42.0	241	2	US-08-778-275-4	Sequence 4, Appl1
45	39.5	42.0	241	2	US-08-824-996-8	Sequence 8, Appl1

ALIGNMENTS

```
RESULT 1
US-09-146-580-3
Sequence 3, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papsidero, Lawrence D
APPLICANT: Dyster, Lyn M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
EARLIER FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 18
TYPE: PRT
ORGANISM: Homo sapiens
US-09-146-580-3

Query Match      100.0%; Score 94; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.7e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 TEVSHHSIRRLRYVMNC 18
Db      1 TEVSHHSIRRLRYVMNC 18

RESULT 2
US-09-146-580-2
Sequence 2, Application US/09146580A
Patent No. 6306653
GENERAL INFORMATION:
APPLICANT: Papsidero, Lawrence D
APPLICANT: Dyster, Lyn M
TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
FILE REFERENCE: 200755/1002
CURRENT APPLICATION NUMBER: US/09/146,580A
EARLIER FILING DATE: 1998-09-03
EARLIER APPLICATION NUMBER: 60/071,889
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 60/092,155
EARLIER FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 18
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (47)
; OTHER INFORMATION: Xaa at position 47 is either Arg or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (68)
; OTHER INFORMATION: Xaa at position 68 is either Lys or Asn
; US-09-146-580-2

Query Match          100.0%; Score 94; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 6.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TEVSHHSRRLERVMNC 18
DB 9 TEVSHHSRRLERVMNC 26

RESULT 3
US-09-146-580-1
; Sequence 1, Application US/09146580A
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (70)
; OTHER INFORMATION: Xaa at position 70 is either Arg or Gly
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa at position 91 is either Lys or Asn
; US-09-146-580-1

Query Match          100.0%; Score 94; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TEVSHHSRRLERVMNC 18
DB 32 TEVSHHSRRLERVMNC 49

RESULT 4
US-09-049-671-1
; Sequence 1, Application US/09049671
; Patent No. 5928874
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Guegler, Karl J.
```

```

; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: NEKI-RELATED PROTEIN KINASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/049,671
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0492 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 293 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRSNOR01
; CLONE: 3069734
; US-09-049-671-1

Query Match          42.6%; Score 40; DB 2; Length 293;
Best Local Similarity 53.3%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHHSRRLERVMNC 18
DB 251 SDHYSEELRLQIVMNC 265

RESULT 5
US-09-295-068-1
; Sequence 1, Application US/09295068
; Patent No. 6030801
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: NEKI-RELATED PROTEIN KINASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/09/295,068
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,671
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0492 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 293 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: UTRSNOR01
CLONE: 3069734
US-09-295-068-1

Query Match 42.6%; Score 40; DB 3; Length 293;
Best Local Similarity 53.3%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 SHHISRLLERYVMC 18
| | | | | : | | | | |
DB 251 SDHYSEELRLQVVMC 265

RESULT 6
US-09-221-235-2
Sequence 2, Application US/09221235
Patent No. 6043040
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,235
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-235-2

Query Match 42.6%; Score 40; DB 3; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 SHHISRLLERYVMC 18
| | | | | : | | | | |
DB 260 SDHYSEELRLQVVMC 274

RESULT 7
US-09-221-928-2
Sequence 2, Application US/09221928
Patent No. 6121030
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115

EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-928-2

Query Match 42.6%; Score 40; DB 3; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 SHHISRLLERYVMC 18
| | | | | : | | | | |
DB 260 SDHYSEELRLQVVMC 274

RESULT 8
US-09-221-527-2
Sequence 2, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-527-2

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 SHHISRLLERYVMC 18
| | | | | : | | | | |
DB 260 SDHYSEELRLQVVMC 274

RESULT 9
US-09-221-236-2
Sequence 2, Application US/09221236
Patent No. 6146841
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MNI-050
CURRENT APPLICATION NUMBER: US/09/221,236
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 2
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-236-2

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 4 SHHISRLLERYVMC 18
| | | | | : | | | | |

Db 260 SDHYSEELROLVVMC 274

```
RESULT 10
US-09-221-416-2
; Sequence 2, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-221-416-2
```

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SHHISRLRLERVVMC 18
| | | | : | | | |
Db 260 SDHYSEELROLVVMC 274

```
RESULT 11
US-09-221-245-2
; Sequence 2, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-221-245-2
```

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SHHISRLRLERVVMC 18
| | | | : | | | |
Db 260 SDHYSEELROLVVMC 274

```
RESULT 12
US-09-163-115-2
; Sequence 2, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
```

```
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-163-115-2
```

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SHHISRLRLERVVMC 18
| | | | : | | | |
Db 260 SDHYSEELROLVVMC 274

```
RESULT 13
US-09-221-528-2
; Sequence 2, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-221-528-2
```

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SHHISRLRLERVVMC 18
| | | | : | | | |
Db 260 SDHYSEELROLVVMC 274

```
RESULT 14
US-09-593-553-2
; Sequence 2, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNT-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; EARLIER FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-593-553-2
```

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SHHISRLRLERVVMC 18
| | | | : | | | |
Db 260 SDHYSEELROLVVMC 274

RESULT 15
US-09-221-237-2
: Sequence 2, Application US/09221237
: Patent NO. 6214597
: GENERAL INFORMATION:
: APPLICANT: Acton, Susan
: TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
: FILE REFERENCE: MNI-050
: CURRENT APPLICATION NUMBER: US/09/221,237
: CURRENT FILING DATE: 1998-12-28
: EARLIER APPLICATION NUMBER: 09/163,115
: EARLIER FILING DATE: 1998-09-29
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 302
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-221-237-2

Query Match 42.6%; Score 40; DB 4; Length 302;
Best Local Similarity 53.3%; Pred. No. 29;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 SHHISRLLERYVMC 18
| | | | : ||||
Db 260 SDHYSEELROLVVMC 274

Search completed: January 14, 2003, 18:17:04
Job time : 5 secs

THIS PAGE BLANK (USPTO)